

A Thesis

X-RAY CRYSTAL ANALYSIS OF ACID SALTS,
AND TRIS-ETHYL-SULPHONYL METHANE.
DEVELOPMENT OF RELATED COMPUTER PROGRAMS

Submitted to the University of Glasgow
for the degree of Doctor of Philosophy
in the Faculty of Science

by

Douglas R. McGregor, B.Sc.

Chemistry Department

June, 1967

ProQuest Number: 11011827

All rights reserved

INFORMATION TO ALL USERS

The quality of this reproduction is dependent upon the quality of the copy submitted.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if material had to be removed, a note will indicate the deletion.



ProQuest 11011827

Published by ProQuest LLC (2018). Copyright of the Dissertation is held by the Author.

All rights reserved.

This work is protected against unauthorized copying under Title 17, United States Code
Microform Edition © ProQuest LLC.

ProQuest LLC.
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106 – 1346

ACKNOWLEDGEMENTS

I wish to express thanks to my supervisors Professor J. M. Robertson and Dr. J. C. Speakman for their encouragement and advice during the course of my work.

I am grateful to Professor D. W. J. Cruickshank, Drs. W. S. MacDonald, W. Oberhansli, J. S. Rollett, J. G. Sime and K. W. Muir, and Messrs. J. G. F. Smith and R. Pollard who have made their computer programs available to me.

I am indebted to Dr. J. C. Speakman for crystalline samples of the various chemical compounds studied, and to Dr. H. H. Mills and Dr. Speakman for their critical and helpful reading of this thesis.

My thanks are due to the Department of Scientific and Industrial Research for their financial support during the first two years of the work.

For many helpful and encouraging discussions during the last three years I am grateful to all my colleagues in the Chemistry Department, and especially to Professor D. W. J. Cruickshank and Dr. K. W. Muir.

I must also thank my wife for her constant help and encouragement throughout.

I would like to express my gratitude to Mrs. J. Crossan who typed the manuscript.

SUMMARY

The phase problem of X-ray crystallography is stated, some of the methods available for its solution are surveyed and their theoretical background is outlined.

Three computer programs written by the author, are described. These are a program for scanning three-dimensional density maps, which automatically produces interpolated coordinates for every peak in the map; a program for sorting crystallographic reflexion data into order by the Miller indices; and a program which supplies a suitable weighting-scheme for structure-factor-least-squares refinement. The "ASS" system of crystallographic computer programs is then described with special reference to the method of storing the crystallographic data, and the features which enable it, in some cases, to obtain the positions of all the light atoms automatically, given the position of only the heavy-atom.

The crystal structure analysis of tris (ethyl sulphonyl) methane is described, and the structure obtained is compared to that of tris (methyl sulphonyl) methane. A description is given of the course of the structure analyses of three acid salts, potassium hydrogen dianisate, potassium hydrogen dicrotonate and rubidium hydrogen dicrotonate, and their structures, which all contain short hydrogen bonds, are discussed.

TABLE OF CONTENTS

	PAGE
CHAPTER 1	
THE PHASE PROBLEM AND ITS SOLUTION	
1.1. INTRODUCTION	1
1.2. THE HEAVY-ATOM METHOD	3
1.3. THE ISOMORPHOUS-REPLACEMENT METHOD	6
1.4. THE ANOMALOUS DISPERSION METHOD	9
1.5. VECTOR SPACE METHODS	13
1.5.1. THE PATTERSON FUNCTION	13
1.5.2. HEAVY ATOM SUPERPOSITION METHODS	14
1.5.3. EQUAL ATOM STRUCTURES	15
1.6. DIRECT METHODS	18
1.6.1. INTRODUCTION	18
1.6.2. INEQUALITIES	18
1.6.3. PHASE-DETERMINING EQUATIONS	22
1.6.4. STATISTICAL METHODS	25
1.7. CONCLUSION	26

CHAPTER 2

CRYSTALLOGRAPHIC COMPUTER PROGRAMS

2.1.	INTRODUCTION	28
2.2.	THE FOURIER SEARCH PROGRAM	29
2.3.	THE SORTING PROGRAM	33
2.4.	THE WEIGHTING-SCHEME PROGRAM	37
2.5.	THE ASS SYSTEM	
2.5.1.	INTRODUCTION	40
2.5.2.	DESCRIPTION	41
2.5.3.	AUTOMATIC STRUCTURE SOLUTION	44
2.5.4.	RESULTS	46

CHAPTER 3

THE STRUCTURE DETERMINATION OF TRIS-ETHYL-SULPHONYL METHANE

3.1.	INTRODUCTION	49
3.2.	EXPERIMENTAL	52
3.3.	COURSE OF STRUCTURE ANALYSIS	58
3.4.	LEAST-SQUARES REFINEMENT	61
3.5.	DIFFERENCE MAP AND LOCATION OF HYDROGEN ATOMS	63
3.6.	POLAR COORDINATE SHIFT CAUSED BY UNCORRECTED ANOMALOUS DISPERSION	64
3.7.	DESCRIPTION AND DISCUSSION OF STRUCTURE OF TESM	67
3.8.	TESM AND TMSM	70
3.9.	APPENDIX: DETAILED COMPARISON OF TESM AND TMSM	72

CHAPTER 4

THE STRUCTURE DETERMINATION OF POTASSIUM HYDROGEN DIANISATE

4.1.	INTRODUCTION	74
4.2.	EXPERIMENTAL	76
4.3.	COURSE OF THE ANALYSIS	76
4.4.	LEAST-SQUARES REFINEMENT	79
4.5.	DISCUSSION OF STRUCTURE	82

CHAPTER 5

THE STRUCTURE DETERMINATION OF POTASSIUM HYDROGEN DICROTONATE

5.1.	INTRODUCTION	85
5.2.	EXPERIMENTAL	87
5.3.	CHOICE OF SPACE GROUPS	88
5.4.	STRUCTURE SOLUTION	89
5.5.	LEAST-SQUARES REFINEMENT	91
5.6.	DISCUSSION	94
5.7.	APPENDIX: CELL DIMENSIONS AND REDUCED CELL	96

CHAPTER 6

THE CRYSTAL STRUCTURE DETERMINATION OF RUBIDIUM HYDROGEN DICROTONATE

6.1.	INTRODUCTION	97
6.2.	EXPERIMENTAL	99
6.3.	STRUCTURE SOLUTION	101
6.4.	DISCUSSION	106

REFERENCES	111
APPENDIX 1	
A.1.1. THE FOURIER SEARCH PROGRAM	116
A.1.2. THE SORTING PROGRAM	125
A.1.3. THE WEIGHTING--SCHEME PROGRAM	130
APPENDIX 2	
ACCURACY OF STRUCTURES	137

CHAPTER 1

THE PHASE PROBLEM AND ITS SOLUTION

Contents

- 1.1. Introduction
- 1.2. The Heavy-Atom Method
- 1.3. The Isomorphous-Replacement Method
- 1.4. The Anomalous Dispersion Method
- 1.5. Vector Methods
 - 1.5.1. The Patterson Function
 - 1.5.2. Heavy Atom Superposition Methods
 - 1.5.3. The Equal Atom Structures
- 1.6. Direct Methods
 - 1.6.1. Introduction
 - 1.6.2. Inequalities
 - 1.6.3. Phase-determining Equations
 - 1.6.4. Statistical Methods
- 1.7. Conclusion

Chapter 1

1.1. Introduction

The central problem in X-ray crystallography is to determine the relative phases of the waves of radiation diffracted by the crystal. Both the phases and the amplitudes must be available in order to combine the waves and to obtain an image of the diffracting crystal structure. While the amplitudes are readily obtained experimentally by measuring the intensities of the diffracted beams, the phases cannot be measured. In many small compounds the atom positions are determined entirely by the requirements of the space group symmetry. In others, only some, or none, of the atom positions are symmetry-determined. The positions of atoms not found by symmetry must be determined from intensity measurements, and the phase problem appears.

The earliest structure solutions were by trial-and-error, but, except for the smallest structures, the number of trial structures required is astronomical. There are two approaches to the phase problem; one is to use more of the information contained in the experimentally measured amplitudes, the other is to change any difficult problem into one which is simpler to solve. Although both these approaches have produced powerful

methods of structure solution, up till now those of the second kind have been the most successful. These are essentially chemical methods which reduce the many atom problem to a one or two atom problem, followed by an iterative structure improvement and refinement. They are the heavy atom and isomorphous replacement methods, and their extensions, the vector superposition and the anomalous scattering method. Of all these the most widely applied is the heavy-atom method.

FIG. 1

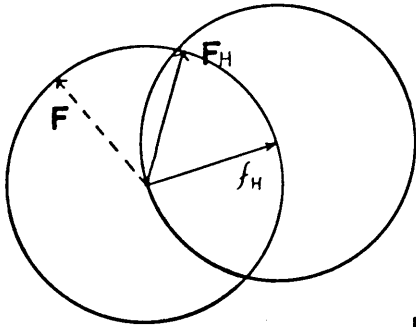


Fig 1a $|f_H| = |F|$

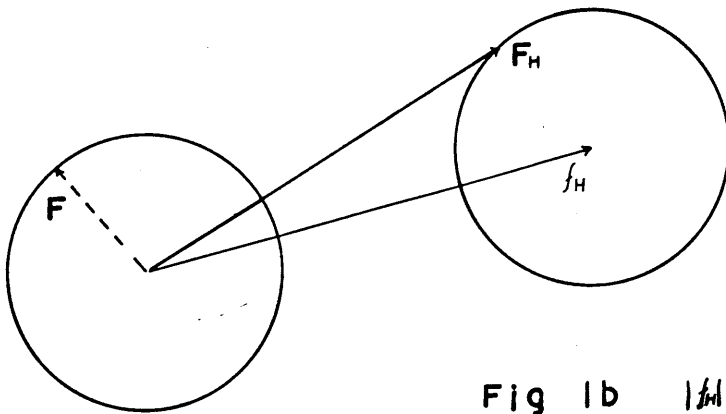


Fig 1b $|f_H| > |F|$

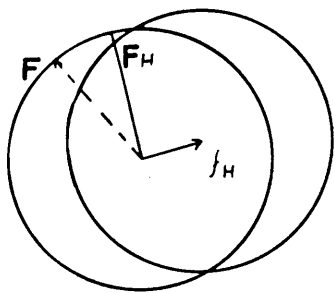


Fig 1c $|f_H| < |F|$

1.2. The Heavy-Atom Method

For this method to be applicable the chemical compound must contain at least one atom which is considerably heavier than the rest, or a heavy atom or group of atoms containing a heavy-atom must be incorporated in the structure by some chemical means. The positions of these heavy atoms in the unit cell of the crystal can then be determined by considering the vector peaks in the Patterson map. From the known positions of the heavy atoms structure factors can be calculated for a hypothetical structure of the heavy atoms alone.

For any one reflexion this calculation gives the contribution of the heavy atoms to the observed structure amplitude. The limitations on the phase of the structure factor of a non-centrosymmetric structure imposed by the size of the heavy atom contributions are shown in Fig. 1 (Robertson, 1963). F is the contribution of the unknown portion of the structure, f_H is the known contribution of the heavy atom, F_H is the resultant structure factor of the heavy atom derivative.

The phase of f_H is known, while that of F is completely unknown. The phase of the resultant structure factor of the whole structure F_H is constrained so that the end of the vector F_H must lie on the ~~right-hand~~ circle.

The diagrams show that where f_H is large, compared to F , the phase angle of the actual structure factor F_H must be similar to that of the heavy-atom scattering vector. As f_H becomes smaller, the probability tends to $1/2$ that the structure factor will have a phase within $\pm\pi/2$ of the heavy-atom phase. Considering the range of possible phase angles of a typical structure factor, it is surprising that the solution of a non-centrosymmetric structure does not present even more difficulty in practice.

The larger the heavy-atom contribution, the more likely is the phase of the structure factor to be close to that of the heavy-atom contribution. By examining a list of the structure factors, a crystallographer can decide which are safe to be included in the electron-density synthesis. (A small computer program can do this equally well.) Sim (1960) has pointed out that this arbitrary procedure is not completely satisfactory. Either too few unreliable terms may be omitted, or else, if too many are eliminated, the electron-density maps may actually be made worse in spite of the phase errors in these terms.

A more objective procedure is to weight the terms according to the reliability of their phases, as suggested

by Woolfson (1958) and Sim (1960). Even so, except where the heavy-atom is barely heavy enough, unweighted electron-density maps are remarkably successful at revealing the structure.

1.3. The Isomorphous Replacement Method

This is not so widely applicable as the heavy-atom method. At least one heavy-atom derivative, isomorphous with the parent compound, is required. In the non-centrosymmetric case, three isomorphous compounds must be used to obtain a complete solution of the phase problem. The requirement of isomorphism is the chief limitation on the method, since, with small or medium-sized chemical molecules, the addition of a heavy-atom frequently forces the compound to crystallise in a different space group. Compounds such as the phthalocyanines (Robertson, 1936) are rare. Sometimes a series of solvates can be used with great effect, (Cheung, 1966), but solvates frequently form crystals of unsatisfactory composition (Islam, 1966). With exceedingly large molecules such as proteins, however, the addition of heavy-atoms may leave the structure relatively undisturbed. The isomorphous replacement has been brilliantly used by Kendrew (1961) in solving the structure of the protein myoglobin, and by Philips (1965) in the structure determination of lysozyme.

Double Isomorphous Substitution

In the case where three isomorphous crystals are available, compound, compound + Heavy-atom 1, compound + Heavy-atom 2, F , FH_1 , FH_2 are the structure factors for

FIG. 2

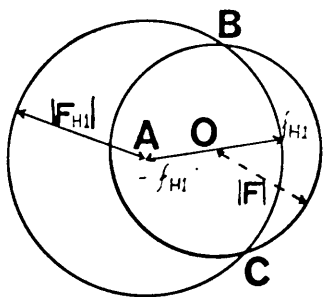


Fig 2a $f_{H1}, |F_{H1}|, |F|$

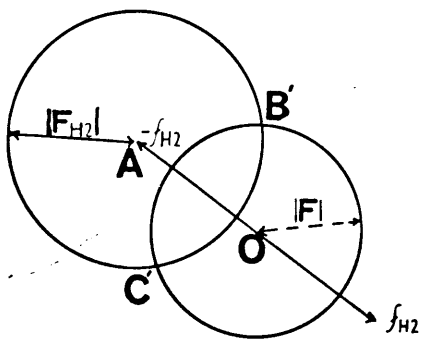


Fig 2b $f_{H2}, |F_{H2}|, |F|$

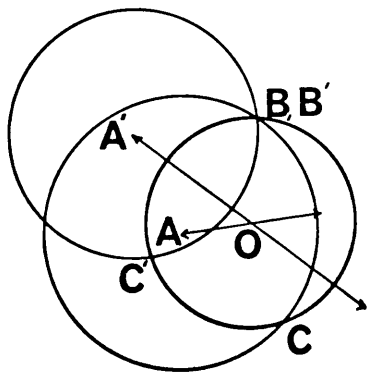


Fig 2c

one reflexion in the three compounds, and fH_1 , fH_2 are the scattering vectors for the heavy-atoms 1 and 2 respectively.

Only the magnitude of F , FH_1 and FH_2 are known, and not their phases. Using vector notation we have:-

$$FH_1 = F + fH_1 \quad (1)$$

and

$$FH_2 = F + fH_2 \quad (2)$$

These can be re-arranged as shown below:-

$$FH_1 - fH_1 = F \quad (3)$$

$$FH_2 - fH_2 = F \quad (4)$$

Since we know both the phase and amplitude of fH_1 and fH_2 , these equations can be solved graphically (Fig. 2) by a procedure introduced by Harker (1956).

From Fig. 2 it can be seen that each isomorphous pair gives rise to two possible solutions. Thus the vector F must be along either OB or OC , and the vector FH_1 must be along either AB or AC . In Fig. 2.b. the vector F must be along OB^1 or OC^1 and the vector F must be along either A^1B^1 or A^1C^1 . When we combine these in a single drawing, Fig. 2c. we can see that there is only one possible solution and the phases of F , FH_1 , FH_2 can be determined unambiguously.

In practice the three circles rarely meet at one

point, and a statistical technique must be used. The "best" intersection-point is usually taken as the centre of gravity of the probability distribution.

1.4. The Heavy-Atom, Anomalous Dispersion Method

This method can be used to obtain more accurate phase-angles, to assist the solution of general non-centrosymmetric structures by the heavy-atom method, to determine absolute molecular configurations directly, and, in some cases, to resolve the problem of "pseudo-symmetry" in the initial heavy-atom-phased electron-density maps.

Pseudo-symmetry occurs whenever the arrangement of the heavy-atoms alone corresponds to a space group more symmetrical than that of the actual crystal. In this case the first electron-density map, phased from the heavy-atom position, is a multiple image of the structure. The weighting procedures discussed earlier do not alter this, and only improve the quality of the multiple image. Usually, structures of this type are solved by selecting some chemically recognisable group of atoms from the initial maps and, after structure factors have been calculated, an improved map is obtained.

Systematic methods are available for non-centrosymmetric structures which make use of the fact that certain atoms scatter X-rays anomalously. They are the direct-phase method of Ramachandran and Raman (1956) and the sine-Patterson method proposed by Okaya, Saito and Pepinsky (1955). Geurtz, Peerdeman and Byvoet (1963) compared the

two methods and concluded that the direct phase method gave the more satisfactory results.

The only disadvantage is that the intensities of both the $F\bar{h}\bar{k}\bar{l}$ and the $Fhkl$ reflexions have to be measured. At present this is laborious, and the method is seldom used. With the coming of computer-controlled diffractometers, data collection will become much more flexible. Since small or accidentally absent amplitudes have little effect on an electron-density map, it would be necessary to phase only the larger terms accurately.

Therefore it would be necessary to measure only the second reflexion $\bar{h}\bar{k}\bar{l}$ if its twin the hkl reflexion were larger than some specified minimum.

The departure from Friedel's Law (1913) is caused by an atom in the structure being excited by the incident radiation. This results in the diffracted wave's having both a real and an imaginary component. The effect is at its maximum if the wavelength of the incident radiation is just longer than the wavelength at which the atom has an absorption-edge, but comparatively small effects can be utilized by the method (Raman, 1958).

F_R, \bar{F}_R are the components of the unknown, normally-scattering atoms;

F_A, \bar{F}_A are the real components of the diffracted beam

FIG. 3

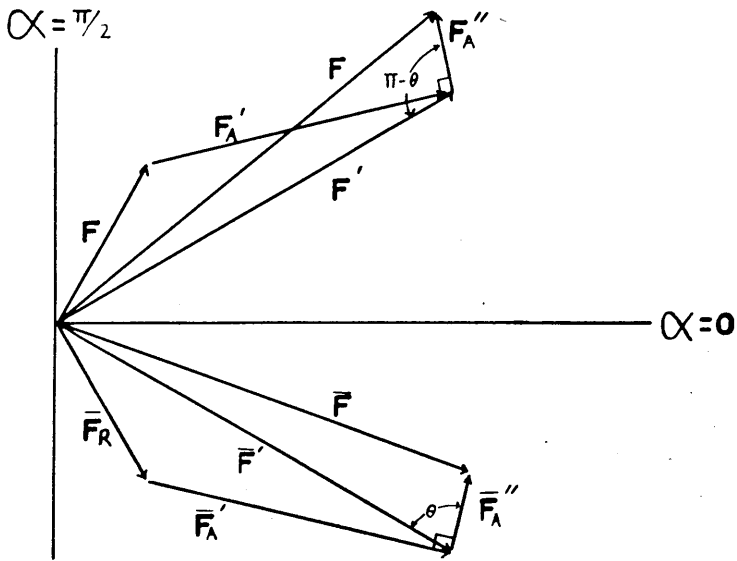


Fig 3a

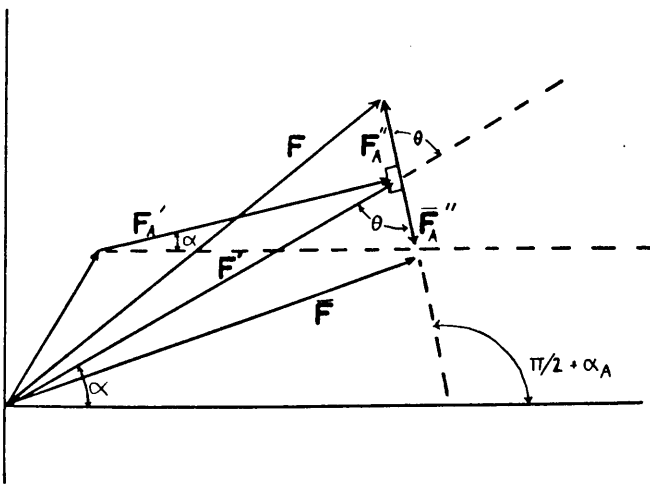


Fig 3b

scattered by the anomalously-scattering atom;

F_A'' , \bar{F}_A'' are the imaginary components of the wave scattered by the anomalously-scattering atom;

F , \bar{F} are the structure factor vectors;

F' , \bar{F}' are the mean structure factor vectors which are to be used in the electron-density map;

α is the required phase angle;

α_A is the known phase angle of the normal component of the heavy atom scattering vector;

θ is the angle between F_A'' and F' .

From the figure it is seen that the phase of F_A is $\pi/2 + \alpha_A$.

Therefore we have

$$\alpha + \theta = \pi/2 + \alpha_A$$

and so:

$$\alpha = \pi/2 + \alpha_A - \theta \quad (1)$$

From the Cosine we have:

$$|F|^2 = |F'|^2 + |F_A''|^2 - 2|F'||F_A''| \cos(180 - \theta) \quad (2)$$

and

$$|F|^2 = |F'|^2 + |F_A''|^2 - 2|F'||F_A''| \cos(\theta) \quad (3)$$

If

$$\Delta|F|^2 = |F|^2 - |\bar{F}|^2$$

$$\text{we have } \cos(\theta) = \frac{\Delta|F|^2}{4|F'||F_A''|} \quad (4)$$

and

$$|F'| = \sqrt{\frac{1}{2}(|F|^2 + |\bar{F}|^2) - |F_A''|^2} \quad (5)$$

The ambiguity in the value of θ obtained from equation (4) is resolved by taking the value of α close to α_A , the phase of the heavy-atom scattering vector. Although the data must be placed on an absolute scale, this may be done adequately (Harris and Mills, 1966) by finding the batch scale factor K from the following expression:

$$K = \frac{\sum^{\text{batch}} |F_c|}{\sum^{\text{batch}} |F_o|}, \quad (6)$$

where $|F_c|$ is the amplitude of the heavy-atom contribution to the structure factor and $|F_o|$ is the experimentally-determined structure amplitude on the relative scale.

Some limitation must be imposed on the values of the shifts from the heavy-atom phase obtained by applying this method. Hall and Maslen (1965) suggest that, if $\cos(\theta) > 1.0$

then $\alpha = \alpha' + \pi/2,$

and if $2|F_A''| > |F|^2 + |\bar{F}|^2,$

then $\alpha = \alpha',$

and if $F' = 0$ or $F_A'' = 0,$

then $\alpha = \alpha'.$

This method therefore seems assured of more frequent use in future.

1.5. Vector Space Methods

1.5.1. The Patterson Function

The first and most important method of displaying the information in the observed structure amplitudes is the Patterson function. This is a Fourier synthesis as shown below:

$$A(uvw) = \frac{1}{V} \sum_h \sum_k \sum_l |F(hkl)|^2 \cos(2\pi(hu + kv + lw)) \dots (1)$$

Patterson derived this synthesis by considering the convolution of the electron-density with itself. Its important property is that the peaks in the Patterson function occur at the ends of the interatomic vectors, for every pair of atoms in the crystal a peak occurs in the Patterson function. Thus there are, in a unit cell containing N atoms, N^2 peaks in the Patterson map and N of these are at the origin.

There are two problems in interpreting the Patterson map:

- (i) Obtaining the actual structure from a set of vector peaks of the structure;
- (ii) Dealing with the lack of resolution of the Patterson which is far from a set of point vectors.

In the matter of obtaining the structure from the Patterson function, the presence of one atom slightly

heavier than the others greatly increases the chances of success.

1.5.2. Heavy-Atom Superposition Methods

One of the first was the vector superposition method of Beavers and Robertson (1950), which utilised the so-called sum function, and was carried out by placing the origin of the Patterson on each heavy-atom position in turn. The positions of the light atoms were then revealed in the sum function. A more popular image-seeking function is the Buerger Minimum Function (Buerger, 1959). In this the value recorded on the resulting map is the lower of the two values at any grid-point, on two displaced Patterson maps.

As carried out by hand, the Minimum Function map does not compare well with a normal heavy-atom phased electron-density map. This may be due partly to the limited accuracy of superposition. In the heavy-atom method it is important that the heavy-atom co-ordinates should be accurate to much better than 0.2 \AA° and interpolation is used to achieve this. By analogy, one would expect that for good results the minimum function must be carried out with interpolated Patterson values. A second point is that the simple minimum-function can be improved if, on choosing one heavy-atom position, other minimum-functions

are obtained for the other heavy-atom positions related by the space group symmetry, and if these minimum-functions are combined, again by the minimum-function.

The advantages of the superposition method are that the accuracy of the image produced does not depend on the weight of the atom, only on its positional accuracy, and on the general accuracy of the observational measurements. Only when the random fluctuations caused by experimental error in the measurements become of comparable size to the image of the structure which is being sought will the result become unreliable.

The disadvantage of the method is that it is much more complicated than the heavy-atom method, and in cases where the heavy-atom is insufficiently heavy to give reliable phasing, it is sometimes difficult to locate the heavy-atom itself.

1.5.3. Equal Atom Structures

Here the problem is to locate one or two atoms, so that the Patterson superposition methods can be used. The most promising approach here is through the S-map, Jacobson, ^{Wunderlich and Lipscomb}(1959), Simpson, Dobrott and Lipscomb: (1965).

The S-Map

This is a special map obtained from the Patterson function by the use of the minimum-function. When an

atom occurs at some position in the unit cell of a crystal, identical atoms must also occur at other positions in the unit cell in accordance with the space-group symmetry. Thus these atoms must give rise to interatomic vectors in the Patterson. The procedure is as follows:- a position in the real space unit cell is selected and, from the space group equivalent positions, the positions in vector space which correspond to these are generated and examined in the Patterson map. The value of $S(x,y,z)$ is the lowest value that occurs at any of the vector positions.

The S-map has the property that it must have density at every point where the structure has electron-density. It is easier to compute than the superposition minimum map, because interpolation is not required, and the map displays all the information that can be obtained directly from the intensities. This is a very powerful function. It has been used to obtain accurate trial atom positions for the superposition method, mainly by Lipscomb and his co-workers. It can also be used to obtain an "M-function" (Woolfson and Main, 1963) from which phases can be determined.

Conclusion

The vector space methods are simpler to understand

than the reciprocal space direct methods, but the large size of the three-dimensional maps makes them difficult to program successfully. They are not rendered powerless by the lack of a centre of symmetry, and some moderately complicated structures have been solved, e.g. Cellobiose, (Jacobson, Wunderlich and Lipscomb, 1959).

1.6. Direct Methods

1.6.1. Introduction

Although the solution of any unknown structure by the heavy-atom method, or any of the methods so far discussed, is a direct solution, the term "direct method" has been restricted by usage to methods which attempt to obtain phase information from the measured amplitudes. The Patterson function displays all the information in the measured amplitudes, but though its form is more understandable, it cannot contain any more "information" than the amplitudes themselves. It is worth noting that, for a thirty-atom problem, a three-dimensional Patterson may contain 50,000 grid-points, whereas there may be only 2,000 intensity measurements, of which perhaps half are reasonably large. Thus the volume of data to be handled is much less for direct methods than for vector methods.

1.6.2. Inequalities

Historically this was the first method. Inequalities arising from different symmetry elements were derived by Harker and Kasper (1949) in their definitive paper.

The theory is based on the application of two well-known inequalities, the Schwarz Inequality and the Cauchy Inequality, to the structure factor equation:

$$F(hkl) = V \int_0^1 \int_0^1 \int_0^1 \rho(x,y,z) e^{2\pi i (hx+ky+lz)} dx dy dz \quad (1)$$

Schwarz's inequality is

$$\left| \int fg dr \right|^2 \leq \left(\int |f|^2 dr \right) \left(\int |g|^2 dr \right) \quad (2)$$

Let us consider a centrosymmetric crystal. If the origin is taken at the centre of symmetry, equation (1) becomes

$$F(hkl) = V \int_0^1 \int_0^1 \int_0^1 \rho(x,y,z) \cos(2\pi (hx+ky+lz)) dx dy dz \quad (3)$$

In equation (2) let

$$\left| \int fg dr \right|^2 \equiv |F(hkl)|^2,$$

$$f \equiv \left(V \rho(x,y,z) \right)^{1/2},$$

$$g \equiv \left(V \rho(x,y,z) \right)^{1/2} \cos(2\pi (hx+ky+lz)),$$

$$dr \equiv dx dy dz,$$

then

$$|F(hkl)|^2 \leq \int_0^1 \int_0^1 \int_0^1 V \rho(x, y, z) dx dy dz \\ \times \int_0^1 \int_0^1 \int_0^1 V \rho(x, y, z) [\cos(2\pi(hx+ky+lz))]^2 dx dy dz \quad (4)$$

Since $\cos^2 \theta = \frac{1}{2}(1 + \cos 2\theta)$, equation (4) can be simplified. The number of scattering electrons (Z) in the unit cell is given by

$$Z \equiv V \int_0^1 \int_0^1 \int_0^1 \rho(x, y, z) dx dy dz \quad (5)$$

Applying these to (4) we get

$$|F(hkl)|^2 \leq Z/2 (Z + F(2h, 2k, 2l)) \quad (6)$$

or

$$\frac{|F(hkl)|^2}{Z} \leq \frac{1}{2} + \frac{1}{2} \frac{F(2h, 2k, 2l)}{Z} \quad (7)$$

By the use of the Cauchy inequality it is possible to produce a stronger inequality, which utilises the fact that atomic scattering decreases in a similar manner for all atoms as the angle of diffraction increases. The inequality obtained is analogous to (7) for the case of a centrosymmetric crystal:

$$\left| U(h, k, l) \right|^2 \equiv \left| \frac{F(h, k, l)}{\sum_{j=1}^N f_j} \right| \leq \frac{1}{2} + \frac{1}{2} \left(\frac{F(2h, 2k, 2l)}{\sum_{j=1}^N f_j} \right) \dots (8)$$

$U(hkl)$ is known as the unitary structure factor, a quantity much used in direct methods.

Conclusion

It is interesting to read that Harker and Kasper saw that the inequalities could be still further strengthened by the use of statistical methods.

"In the case of crystals with large unit cells, it may be possible to divide the values of F_{hkl}/Z by numbers which decrease much more rapidly with increasing $\sin\theta/\lambda$ than do the values of \hat{f} ."

$$\left(\hat{f} \equiv \frac{1}{Z} \sum_{j=1}^N f_j \right).$$

In practice the method has solved very few structures. Its importance is in the impetus it gave to the search for methods of obtaining phases directly from measured

amplitudes.

1.6.3. Phase-Determining Equations

Sayre's Equation

The next advance was made by Sayre (1952) who considered a hypothetical crystal with an electron-density which was everywhere the square of the electron-density in the actual crystal. For the real crystal

$$\rho(x) = \frac{1}{V} \sum_H F_H e^{-2\pi i H x}, \quad (1)$$

while for the "squared crystal",

$${}^{sq}\rho(x) = \rho(x) \times \rho(x) = \frac{1}{V} \sum_H {}^{sq}F_H e^{-2\pi i H x} \quad (2)$$

The structure factors for the squared crystal differ from those of the real crystal. The difference between the structures is that the squared atoms are much "sharper" than the real atoms. Thus the ${}^{sq}F_H$ structure factors are "sharpened" (in the Patterson sense) compared to the F_H series. That is, F_H and ${}^{sq}F_H$ are related by a factor which takes account of this change of shape of the atoms. In general, this factor is different for different values of H ,

$${}^{sq}F_H = g_H F_H. \quad (3)$$

Thus

$$\rho(x)^2 = \rho(x) \times \rho(x) = \left[\frac{1}{V} \sum_H F_H e^{-2\pi i H x} \right]^2 \quad (4)$$

$$= \frac{1}{V} \sum_{H_1} \left[\frac{1}{V} \sum_{H_2} F_{H_1} F_{H_2} \right] e^{-2\pi i (H_1 + H_2) x} \quad (5)$$

$$\text{i.e. } \rho(\chi)^2 = \frac{1}{V} \sum_{H_1} \left[\frac{1}{V} \sum_{H_2} F_{H_1} F_{H_2} \right] e^{-2\pi i (H_1 + H_2) \chi} \quad (5)$$

Let $H_1 + H_2 = H$, and $H_1 = H'$. Then since the summations in (5) run over all integers, the range of H is the same as the range of H' .

Thus

$$\rho(\chi)^2 = \frac{1}{V} \sum_{H'} \left[\frac{1}{V} \sum_{H'} F_{H'} F_{H-H'} \right] e^{-2\pi i H \chi} \quad (6)$$

and comparing (2), (3) and (6), we get

$$\text{sq } F_H = \frac{1}{V} \sum_{H'} F_{H'} F_{H-H'}, \quad (7)$$

or

$$g_H F_H = \frac{1}{V} \sum_{H'} F_{H'} F_{H-H'}, \quad (8)$$

whence

$$F_H = \frac{1}{g_H V} \sum_{H'} F_{H'} F_{H-H'}. \quad (9)$$

This final equation must hold for crystals composed of equal resolved atoms. The equation will hold for any structure factors, including unitary or normalised structure factors. It is also applicable directly to non-centrosymmetric crystals. In this last case F_H is complex.

$$\text{Let } F_H = (A_H + i B_H) = \frac{1}{g_H V} \sum_{H'} (A_{H'} + i B_{H'}) (A_{H-H'} + i B_{H-H'}) \quad (10)$$

$$\text{Then } A_H = \frac{1}{g_H V} \sum_{H'} (A_{H'} A_{H-H'} - B_{H'} B_{H-H'}), \quad (11)$$

$$\text{and } B_H = \frac{1}{g_H V} \sum_{H'} (A_{H'} B_{H-H'} - B_{H'} A_{H-H'}) \quad (12)$$

$$\text{Now } \tan \phi_H = B_H / A_H, \quad (13)$$

$$\text{so } \tan \phi_H = \frac{\sum_{H'} (A_{H'} B_{H-H'} + B_{H'} A_{H-H'})}{\sum_{H'} (A_{H'} A_{H-H'} - B_{H'} B_{H-H'})}. \quad (14)$$

It can be seen from the form of equations (9) and (14) that they are iteration formulae. They give the values of the structure factor or phase in terms of all the other structure factors. They cannot be applied directly until at least some signs or phases have been determined. They have, however, been used to improve sets of signs to make them self-consistent. Equation (9) has been used by Cochran and Douglas (1957) to evaluate various sets of signs obtained by another method. This was done by calculating the value of an index k for each set of signs, where

$$k = \frac{\sum_{H'} |F_H - \frac{1}{g_H \sqrt{V}} \sum_{H'} F_{H'} F_{H-H'}|}{\sum_{H'} |F_H|}. \quad (15)$$

This has a value close to zero for the correct set of signs.

The initial sign determination is often accomplished by means of a degenerate form of Sayre's equation. If, in equation (9), one of the terms in the summation is very large, the sign of the total may be taken as the sign of this largest term,

$$S_H \approx S_{H'} S_{H-H'}, \quad (16)$$

and if two or more terms in the summation are large, and the structure factor on the left hand side is also large, it is taken that the large terms all have the same sign.

1.6.4. Statistical Methods

This method stems from the work by Hauptman and Karle (1952) (1953). These workers developed mathematical relationships involving normalised structure factors. This is defined as

$$E = \frac{1}{\langle \overline{F^2} \rangle}^{1/2} F$$

where E is the normalised structure factor and

$$\langle \overline{F^2} \rangle$$

is the "local average" value of F^2 .

The magnitudes of the E values are thus independent of the number of electrons in the structure.

Even so, their undoubted success seems to stem mainly from their meticulous attention to practical details rather than from any of the special formulae used.

Practical points which must not be overlooked are:-

1. The collection of data for as much of reciprocal space as possible.
2. The data must be scaled, to obtain a single relative scale.
3. Great care must be taken in obtaining the value of

$\langle F^2 \rangle$ to transform the structure amplitudes to normalised amplitudes.

4. The use of a method, the "Symbolic Addition Method", which assigns letter phases to various large E values. The technique is based entirely on the Karle and Hauptman Σ_2 relationship, which is very similar to Sayre's equation.
5. The initial set of signs is made self-consistent, and is expanded by the use of the Σ_2 equation.
6. If any signs are still in doubt, electron-density maps are calculated for all possible sets of signs. The co-efficients used in this Fourier synthesis are the normalised structure factors, and the resulting map is a kind of "sharpened" electron-density map.

By using this method the Karles have solved many crystal structures (1957) (1961) (1963) (1966A, B, C), including a non-centrosymmetric crystal structure of some complexity

1.7. Conclusion

It seems certain that in the future crystal structures of compounds containing only light atoms will be solved more frequently , especially if the substance studied is precious, or the attachment of a suitable heavy-atom

proves difficult. Direct methods, vector methods and anomalous dispersion methods may also be used to supplement the basic heavy-atom and isomorphous replacement techniques.

"It may be better to spend a few days in the calculating room than six months in the laboratory".

CRYSTALLOGRAPHIC COMPUTER PROGRAMS

Contents

- 2.1. Introduction
- 2.2. The Fourier Search Program
- 2.3. The Sorting Program
- 2.4. The Weighting-scheme Program
- 2.5. The ASS System
 - 2.5.1. Introduction
 - 2.5.2. Description
 - 2.5.3. Automatic Structure Solution
 - 2.5.4. Results

Chapter 2

CRYSTALLOGRAPHIC COMPUTER PROGRAMMING

2.1. Introduction

While the work for this thesis was being done an English Electric KDF 9 computer was installed at Glasgow University. This replaced the older DEUCE computer which had been used for all the crystallographic calculations, and the changeover necessitated the re-writing of all the computer programs.

When a fast modern computer became available, it was apparent that the bottleneck of a crystal structure analysis was the stage at which the crystallographer plotted out his maps, examined the results of his labours and decided what to do next. It therefore seemed desirable to shift some of the burden of the interpretative work from the crystallographer to the computer.

The programs described in this chapter are the author's contributions to this.

2.2. The Fourier Search Program

Description

In structure analysis the interpretation of large, three-dimensional maps of all kinds is a very important job. The program described below can help the crystallographer in this.

The program scans any three-dimensional map and picks up the approximate positions of all the maxima above a given threshold value. From these approximate positions and the density values surrounding them the program then calculates the interpolated co-ordinates of the maximum, the maximum density value (or peak height) and the integrated density (or peak weight). The program then sorts the peaks into order of decreasing density and outputs them.

The map to be searched must be input to the program in an agreed format from a magnetic tape. It has been used successfully to search Patterson maps and "difference" maps as well as the more usual electron-density maps. A second version is available which searches for minima, and this has been used successfully to search for hydrogen peaks in the maps produced in structure analysis by neutron diffraction.

Method

The program requires as input a magnetic tape on which the map to be scanned is stored in a binary form. This is usually calculated by the Glasgow Fourier program (Dr. J. G. Sime).

The map is searched, one section at a time, and a list is obtained of all the maxima greater than the threshold value which the user has specified. This list of peaks in any one section is then examined to determine whether a higher maximum for any peak occurs in another section. A second check is made on whether it is possible that the maximum of any peak lies outside the computed section of the map. If either of these occurs, then the peak is ignored. In this way, the approximate co-ordinates of each distinct maximum are obtained.

A second search through the map is then made, and the density values of the 27 grid points containing each maximum are picked from the map.

The interpolation follows that of Shoemaker, Donohue, Schomaker and Corey(1950). An expression

$$\begin{aligned} \rho(x,y,z) = & Ax^2 + By^2 + Cz^2 \\ & + Dx + Ey + Fz \\ & + Gxy + Hyz + Ixz \\ & + J \dots\dots\dots(I), \end{aligned}$$

is used, and the coefficients are fitted to the density values by the Least-Squares Method. Originally the logarithm of the density was fitted. Since atomic electron-density is Gaussian, this was expected to be the best method. When an electron-density map is not well-phased, or if it suffers from severe termination-of-series-error, the peaks are often truncated or misshapen. It was found by experience that (I) gave a much better fit to the interpolated maximum peak height. No doubt the logarithmic formula could have been made to fit better by using a suitable weighting scheme in the Least-Squares calculation, but (I) has been found to give excellent results in practice.

The co-ordinates of the maximum are then found by evaluating the partial derivatives of (I) with respect to x , y and z , equating these to zero and solving for the co-ordinates x , y and z . Two difficulties may arise. The first is that the expression for the density fitted to the maximum may not tend to zero in all directions. This results in a failure of the solving procedure. The second is the related difficulty that the "interpolated" maximum may appear to be outside the block of 27 grid-points. This is physically unreasonable. Should either of these difficulties arise the

program outputs the approximate co-ordinates of the peak, with a marker, at the end of the main list of maxima.

The program does a simple numerical integration over the block of 27 grid-points surrounding each peak. Since this volume is less than the total volume of a typical peak, the result is always considerably smaller than the chemically expected number of electrons. The volume of integration is kept small because density maps often lack resolution and, on integration over a large volume, would include density from neighbouring peaks. The integrated density is a better criterion of the reliability of a peak than the maximum height, because this height also depends on the vibration amplitude of the atom concerned. The list of peaks is therefore sorted into descending order by peak-density.

Results

The program has been in regular use in Glasgow and at several other universities for over a year. It has been found to be very reliable and to produce accurate co-ordinates. It is used for searching all types of maps and several structures have been solved without a single map being plotted out.

2.3. The Sorting Program

Introduction

Since the actual operation of sorting is easy to understand, and mechanical card sorters work very efficiently, the difficulty of sorting large quantities of crystallographic data, stored on paper tape or magnetic tape is not, at first, obvious. The problem becomes serious only when the information to be sorted exceeds the fast-storage capacity of the computer.

Sorting Methods

Two methods suggest themselves. The first is the sort-and-merge technique. From information theory this can be shown to be, in general, the most efficient technique. It requires at least three magnetic tapes, one for input and output of data and two for temporary storage of partially sorted data. As much information as possible is read into the core-store of the computer, where it is sorted, and output to one of the two "work" tapes. Another core-full of data is then read in, sorted and output to the other of the two "work" tapes. In this way all the data are given a preliminary sorting. The partially sorted data are then compared, term by term, one term from each of the two batches, to produce ordered sequences of double the size of the original

batches, and this is repeated until finally the sorting is complete.

The second method is to select batches for sorting from the original data. Thus, if the core could hold 1,000 items, the largest (or smallest) thousand would first be selected and sorted. This second method was adopted for this program as it is simpler and requires fewer magnetic tapes.

Description

The program accepts data from a magnetic tape. The format is the one used in the ASS system, but it is inefficient if many passes have to be made through the data. The data are therefore read into the computer and stored on a "work" tape in batches of fifty reflexions. The program also requires the information defining the required order of the data, - which index is going to change most slowly, which most quickly, which at a medium rate and whether the order is to be ascending or descending. For each reflexion plane, therefore, an index number is generated from the crystallographic indices. It is this composite index which will be sorted into order. Its value is given by:

$$\text{Index} = 10,000 \times s + 100 \times m + f,$$

where s is the value of the slowest-changing index;
 m is the value of the medium - changing index;
 f is the value of the fastest-changing index.

A histogram of the number of data with each of the values of the slowest changing index is also constructed. From the histogram maximum and minimum values of the slowest changing index are then chosen so as to define the biggest possible batch which can be stored, complete, in the core store of the computer.

These batches are then selected, sorted and output. The actual sorting is done in two stages:

1. The sorting index and only the address of each reflexion are sorted. The method follows that of Hibbard's program C, (1963). This takes advantage of order in the data, and for unbiased data the time taken is approximately dependent on $n \log_2 n$ where n is the number of items to be sorted.
2. The second stage is to transfer the reflexions to their new addresses.

Finally, the sorted data are output, again in the standard ASS format, on to the output magnetic tape.

Results

The program has been in use in Glasgow for some

time. Typical times required are shown below, (using the unoptimised Kalgol translator):-

No. of Planes	No. of items per Plane	Time
4,400	12	4 minutes
800	14	49 seconds

2.4. Weighting Scheme Program

Introduction

This program will calculate a satisfactory weighting function and will output a paper tape giving an individual weight to each reflexion. The resultant distribution of $\overline{W \Delta^2}$ as a bivariate function of $|F_o|$ and $\sin^2 \theta / \lambda^2$ is also calculated. (The symbols have their usual significance.)

Theory

In a Least-Squares Refinement the weight given to any observation should be $1/\sigma^2$, (Cruickshank, 1964) where σ^2 is the variance of the observation.

In crystal structure analysis, however, it is customary to have only one observation for the majority of the reflexions, thus making a direct estimate of σ^2 impossible. A second difficulty is that a direct estimate of σ^2 from the observations will not include the variance due to systematic errors which may be present in certain classes of observations. The direct estimate of σ^2 may, therefore, be much too small for such classes of reflexions.

Cruickshank shows that a better method of estimating σ^2 , which includes random errors, systematic errors present in the data but not paralleled in the model, and

systematic errors present in the model, but not present in the data, is to take:

$$\sigma^2 = \Delta^2$$

$$\text{where } \Delta = |F_o| - |F_c|$$

Of course, for any individual observation Δ^2 may be a very bad estimate of the variance, but in crystal structure analysis the number of observations is always large. Thus the Cruickshank procedure is to divide the data into batches by $|F_o|$ and $\sin^2\theta/\lambda^2$ and to use the average value of Δ^2 for the batch as a measure of the variance.

Method

The present program divides all the reflexions into batches depending on $|F_o|$ and $\sin^2\theta/\lambda^2$, and the average $\overline{|F_o|}$ and $\overline{\sin^2\theta/\lambda^2}$ and $\overline{\Delta^2}$ are calculated for each batch.

An expression of the form:

$$\log_e(\overline{\Delta^2}) = \{ A + B \overline{F_o} + C \overline{F_o}^2 (+ \dots D \overline{F_o}^3 \dots) + E \lambda^2 / \sin^2\theta + F \lambda^4 / \sin^4\theta + G \sin^2\theta / \lambda^2 \} \dots (1)$$

is used. The coefficients are fitted by the Least-Squares Method. The observational equations are weighted in proportion to the numbers in each batch. Very small batches are ignored and slightly larger batches are tested to ensure that the value of $\overline{\Delta^2}$ is not dispropor-

tionately small. Isolated reflexions with large unrepresentative values of Δ^2 are also omitted from the calculation of the weighting expression.

From the expression:

$$W = \exp - \left\{ A + B|F_0| + C|F_0|^2 (+ \dots D|F_0|^3 \dots) + E \lambda^2 / \sin^2 \theta + F \lambda^4 / \sin^4 \theta + G \sin^2 \theta / \lambda^2 \right\} \dots (2)$$

weights are now calculated for every reflexion. Provision is also made for scaling the weights to a suitable relative scale, if this is desired. The output tape is in the format required for the Cruickshank Structure Factor Least-Squares Program.

Results

This is one of the most recent programs so there has not yet been wide general usage, but the structures described in this thesis were all weighted by it. The distribution of $W \Delta^2$ which it produces are generally satisfactory and the time taken is only a few seconds for the average structure.

2.5. The ASS System

2.5.1. Introduction

This was set up shortly after the completion of the Fourier Search program had shown that the computer interpretation of electron-density maps was possible. Its name meant Automatic Structure Solution but it was also intended to convey that this was to be a simple approach and that it should not be left to do too much unattended!

The main feature is that ASS can carry out cycles of calculations, - structure factor calculation, Fourier summation, Fourier search, structure factor calculation, Fourier summation etc. - with very little attention from the crystallographer.

A second important feature is that the system of storing crystallographic data on magnetic tape can be used easily by inexperienced programmers. In fact, the system was deliberately kept simple, at the cost of complete generality, for this reason.

A third feature is that the handling of data and intermediate results on magnetic tape has largely eliminated the punching of very large quantities of paper tape, which was time consuming and much less reliable than has been the case with magnetic tape.

The work was carried out in collaboration with

Mr. K. W. Muir and Mr. R. Pollard, who discussed the design with the author and who have both written programs for the system.

Several other programs have been adapted, principally J. G. Sime's Fourier Program, and various other people have contributed programs.

2.5.2. Description

Data Handling

In crystal structure analysis, there are two types of information to be stored.

Type 1

This may be thought of as "permanent" information. It is already possessed by the crystallographer at the start of the attempt to solve the structure, and is seldom changed in the course of the refinement. Examples are:-

- unit cell dimensions,
- space group symmetry,
- chemical elements present and their
atomic scattering factors, and
- experimental reflexion data.

Type 2

This information relates to the model proposed for the structure, and changes as the solution and refinement proceed. Examples are:

atomic co-ordinates and other parameters,
structure factors,
modified structure factors,
maps, and
peak co-ordinates.

All this information is stored on magnetic tape which acts as a "file".

Since this is a simplified data-storage system, only one set of "temporary" type 2 data is kept in the file. The temporary data are over-written when the next iteration commences.

The permanent data are loaded from paper tape. This is done only once per structure when computation for the structure analysis is started. The current atomic parameters, also on paper tape, are added to the file once per cycle, to replace the previous atomic parameters. Peak co-ordinates from maps are stored directly on the magnetic tape by the Fourier Search Program, which also prints them, and punches them out on paper tape.

Should an accident happen to the magnetic tape file, - failures are fortunately a rare occurrence, all the information to load a new magnetic file tape is thus available on paper tape, although structure factors may have to be re-calculated.

Crystallographic information has a very varied format. One approach to this problem is to use a system of "keys", as in the "Index of Lists" compiled by Cruickshank, Freeman, Rollett, Sime, Smith, Truter and Wells (1964). The approach in the ASS system is to define rigorously the format of the information which is stored. For example, it was decided that every structure factor array would be named "SFS" and that its first three items would always be the crystallographic Miller indices of the structure factor; the fourth and fifth items would be $|F_o|$ and $|F_c|$ respectively, and so on for all the other items (McGregor, 1966). This is not such a general method as the "keys" method, but since the programmers were all working together, lack of generality was not a serious disadvantage and, moreover, it simplified the programs considerably.

The variable sizes of the blocks of information are dealt with in two ways:

(A) For reflexion data, structure factors and modified structure factors the tape has one block per reflexion. The last block contains a "last block marker". No check is kept on the number of individual blocks which can therefore easily be altered at will by adding or deleting individual reflexions.

(B) In front of individual blocks of variable size are small "key" blocks of fixed size. Each small block gives the size of the variable blocks following it.

Programs for ASS

Since this is basically a data-handling system, many ASS programs have no connection with "Automatic Structure Solution", e.g. R. Pollard's Isotropic Structure-Factor - Least-Squares Program. These will not be discussed here.

2.5.3. Automatic Structure Solution

When the Fourier Search Program became available it was seen that automatic Fourier refinement and structure solution were at least possibilities. There are two difficulties.

Not all the peaks on an electron-density map correspond to genuine atoms. They may be diffraction ripples or peaks caused by incorrect phasing of the structure amplitudes. When the map is phased by structure factors calculated from a heavy atom in a special position the peaks may, in fact, be multiple images of the light atoms in the structure. The second difficulty is that, once an incorrect peak is accepted as an atom and is included in the phasing calculation, it will appear in the next

electron-density map. This is particularly true of non-centrosymmetric refinements. An automatic Fourier refinement system must therefore apply a series of tests to eliminate the spurious peaks picked up by the Fourier Search Program.

A crystallographer can look at a three-dimensional plot of electron-density or peak positions and decide intuitively which peaks agree with what is known of the chemical structure and whether the packing of the molecules is feasible. A program to do this would need to be fairly sophisticated. Kitaigorodskii (1965) has suggested that the criterion of packing may be sufficient to determine organic structures.

Another method, which has been applied by J. S. Rollett (1965) is to carry out one or two cycles of Least-Squares Refinement, refining the occupation number of the "unknown" peaks.

The ASS testing scheme for peaks is somewhat simpler than either of the other methods, and it is incorporated in the Structure-Factor-with-Testing Program. This was written in collaboration with Mr. K. W. Muir. The peaks are fed, one at a time, into a structure factor calculation over a randomly-selected sub-set of the data. The contributions of the tentative atom are combined with the

contributions of the known atoms. Those tentative atoms which produce a specified drop in the R-factor are accepted. Where pseudo-symmetry is present the peak is tested in all the possible positions and the position giving the lowest R-factor is adopted. After this preliminary test the tentative atoms which were rejected are re-tested to determine whether, in the company of the other newly accepted atoms, they now lower the R-factor. The iterative process is continued until no more tentative atoms are accepted.

The newly accepted atoms are now subjected to a second test. They are re-tested to determine whether they still lower the R-factor, and, in the cases where pseudo-symmetry was present, whether the atom is in the best position. This second test is re-cycled until no further changes occur.

2.5.4. Results

The system has now been in operation at Glasgow for about ten months. During this period it has been used for the refinement of almost all the crystal structures, up to the point at which a transition has to be made to refinement by an Anisotropic Least-Squares Program.

In spite of this, few attempts have been made to solve structures automatically. Most crystallographers

seem to have an instinctive distrust of the automatic procedure. They prefer to make a model, or plot the peak positions. Having decided how the molecule lies, they often check that the peaks which they have chosen do lower the R-factor. Altogether, about twenty-five structures have been solved using the system in this way.

Nevertheless, some examples of automatic structure solution can be cited.

Potassium Hydrogen Di-Crotonate

This structure had already been solved by conventional methods and is described elsewhere in this thesis. The space group is $P\bar{1}$ and the formula is $KC_8O_4H_{11}$. Starting with only the potassium atom, after two cycles, eleven out of the thirteen atoms, excluding hydrogen, had been picked out correctly. The two missing atoms were carboxyl carbon atoms which had been dropped by the Fourier Search Program because the interpolation matrix was singular. No incorrect peak was accepted as an atom although the lowest correct atom was twenty-second on the peak list. (This structure gave electron-density maps with many diffraction ripples because only three layers of reflexion data were collected up the b-axis.)

Thus the gross structure of this compound was obtained after an expenditure of less than 60 minutes' computer time.

Tri-Nitro-Bromo fluorenone - Anthracene Complex

(Solved by Mr. R. Pollard and Dr. G. Ferguson)

The space group is $P2_1/C$ and the chemical formula is $C_{34}(NO_2)_3Br$. The R-factor for the bromine atom alone was 58%. After two cycles the expected number of atoms had been picked out and the R-factor had been reduced to 36%. On examining the list of peak heights it was decided that disorder was present and that the tri-nitro-bromo-fluorenone molecule could pack in alternative ways. This meant that the bromine atom and one of the nitro-groups were disordered. Without altering the structure in any way except to adjust the scattering factors of the bromine and nitrogen atoms involved, the R-factor fell immediately to 29%. The refinement has since been continued to a terminal R-factor of 8%.

Tris-Ethyl Sulphonyl Methane

The solution and refinement of this compound will be dealt with elsewhere in this thesis.

THE STRUCTURE DETERMINATION OF TRIS-ETHYL-SULPHONYL
METHANE

Contents

- 3.1. Introduction
- 3.2. Experimental
- 3.3. Course of Structure Analysis
- 3.4. Least-Squares Refinement
- 3.5. Difference Map and Location of Hydrogen Atoms
- 3.6. Polar Coordinate Shift caused by uncorrected anomalous Dispersion.
- 3.7. Description and Discussion of Structure of TESM
- 3.8. TESM and TMSM
- 3.9. Appendix: Detailed Comparison of TESM and TMSM

THE CRYSTAL STRUCTURE. DETERMINATION OF
TRIS-ETHYL-SULPHONYL METHANE $\text{C}(\text{SO}_2\text{CH}_2\text{CH}_3)_3$

3.1. Introduction

Tris-sulphonyl-methane compounds, and their unusual strongly-acidic properties have been known for a long time (Fromm, 1899; Samen, 1941, 1942, 1947). Right from their discovery, speculation has existed as to whether the central carbon atom of the methane molecule lies in the plane of the three sulphur atoms, or whether the molecule has a tetrahedral configuration. If different alkyl groups could be attached to each sulphur atom, the molecule, if non-planar, would rotate the plane of polarised light. Attempts to prepare unsymmetrical tris-alkyl-sulphonyl-methanes and resolve the mixture into its optically-active components all failed. (Gibson, 1931; Kipping, 1935; Bohme and Marx, 1949). In 1955, Doering and Levy suggested that the acidic proton would be lost more readily from a planar than from a non-planar molecule.

Initial X-ray diffraction work on tris-methyl-sulphonyl-methane (TMSM) (Speakman and Abrahams, 1956) revealed a most unusual diffraction effect. In addition to sharply-defined normal X-ray reflexions from which a

rhombohedral space group was deduced, remarkable diffuse reflexions were observed (See Fig. 1). These occurred at reciprocal lattice positions of the hexagonal cell forbidden to the rhombohedral space group. Since the diffuse reflexions remained unaltered in photographs taken at 78°K, they decided that the phenomenon was caused by disorder of the crystal lattice, and not by any molecular vibration effect. Further support was lent to this conclusion by a study of the thermal capacity of TMSM (Staveley and Davies, 1956) in which no phase transitions were observed between 14°K and room temperature.

The crystal structure of the ammonium salt of TMSM was determined by Hoogsteen (1957). Its structure was found to be planar.

Silverton, Gibson and Abrahams (1966) investigated TMSM by three-dimensional X-ray methods. They were unable to refine an ordered structure and found that, to explain anomalous features of their Patterson and difference maps, they had to postulate a disordered structure containing two crystallographically-different molecules, type A and type B, in the ratio of 61:39 respectively. The molecules which are non-planar are stacked in columns, each column having 3-fold symmetry and lying along the 3-fold axes of the space group. Each column contains only one type of

FIG. 1

Weissenberg photograph of $hkio$ reciprocal lattice net
of TMSM, showing diffuse reflexions.

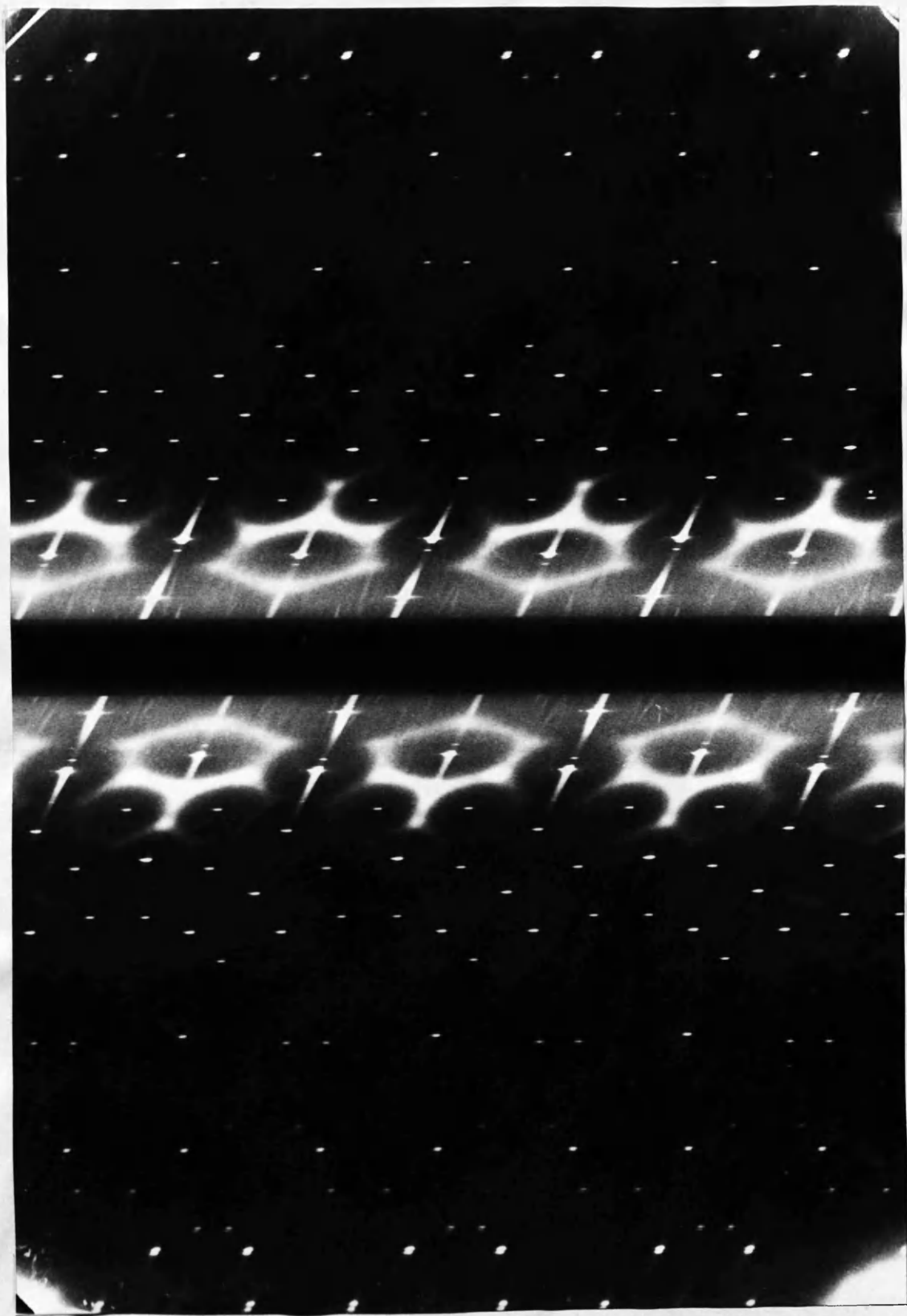
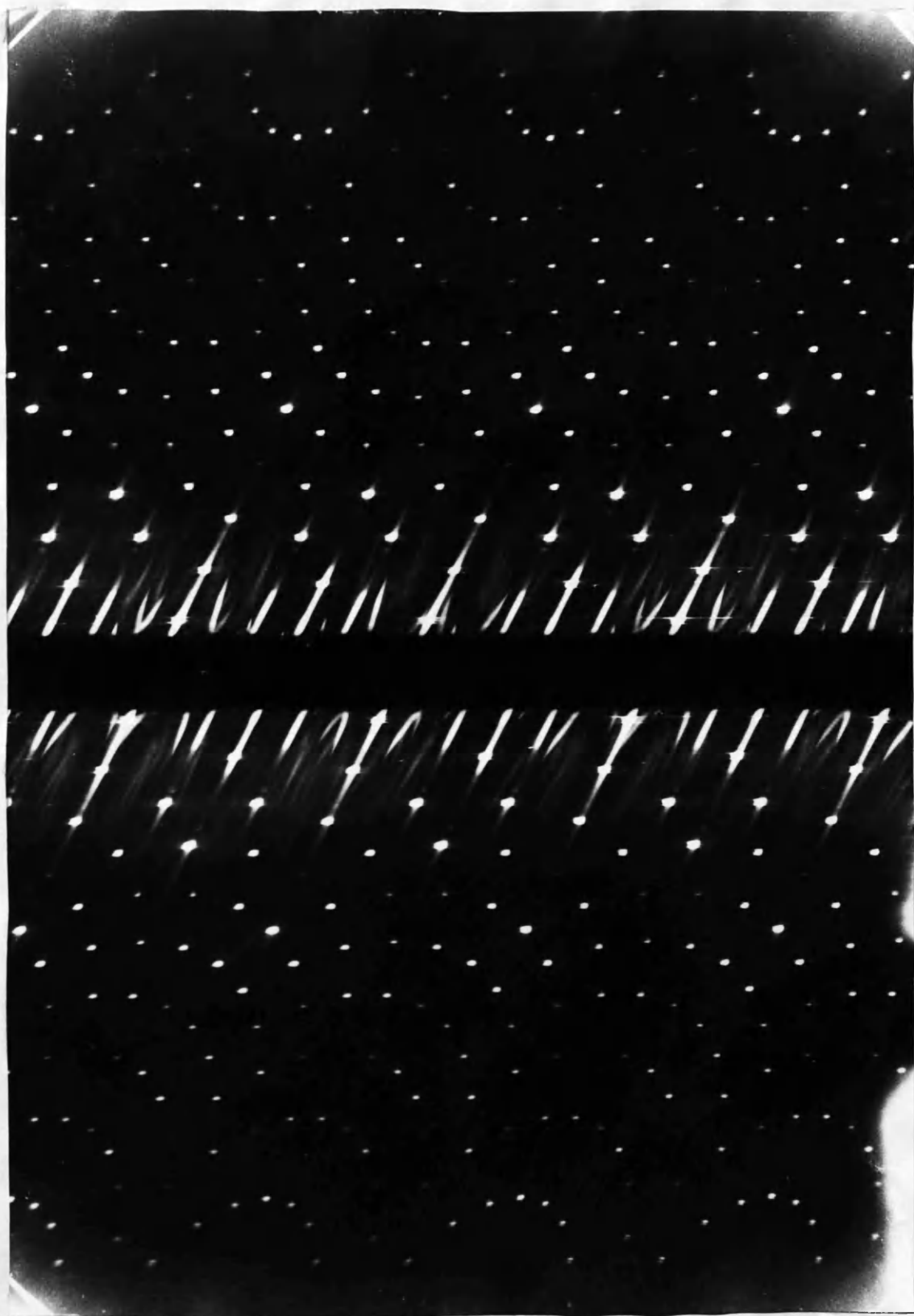


FIG. 2

Weissenberg photograph of h_ki₀ reciprocal lattice net
of TSM.



molecule, and the columns of different types are randomly arranged so that the overall ratio of A:B molecules is maintained. The bond lengths in both type A and type B molecules are chemically-reasonable, with normal intermolecular contacts between stacks of type A and stacks of type B, or between stacks of type A and type A, or between stacks of type B and type B.

Tris-ethyl-sulphonyl methane (TESM) is very similar chemically to TMSM. It belongs to the same rhombohedral space group, but no diffuse X-ray reflexions are observed experimentally (See Fig. 2).

It was therefore thought that this structure would provide a useful comparison with the disordered TMSM. The complete crystal structure has been determined, and the molecular bond lengths have been measured accurately. The crystals used in this determination were prepared by Dr. D. T. Gibson.

3.2. Experimental

Crystal Data

Tris-ethyl-sulphonyl-methane $(\text{CH}_3\text{CH}_2\text{SO}_2)_3\text{CH}$;
molecular weight = 292.40; Laue group $\bar{3}m$;
system rhombohedral;
hexagonal lattice constants $a = b = 14.965$, $c = 9.81\overset{\circ}{\text{A}}$,
 $\gamma = 120.0^\circ$; volume = $1904 \overset{\circ}{\text{A}}^3$; six molecules per unit
cell; $D_m = 1.55 \text{ g cm}^{-3}$; $D_x = 1.565 \text{ g cm}^{-3}$;
 $F_{000} = 942$ electrons;
absorption coefficient for $\text{MoK}\alpha = 5.83 \text{ cm}^{-1}$.

Choice of Space Group

From the Laue Group the space group is either $R\bar{3}c$ or $R\bar{3}c$. Since there are six molecules in the unit cell, and the molecule can have a 3-fold axis, the more obvious choice is $R\bar{3}c$. The alternative is that the molecule would statistically half-occupy the twelve-fold positions in $R\bar{3}c$, although no diffuse reflexions of the type found in TESM were observed.

The initial choice of $R\bar{3}c$ has been justified by the successful solution and refinement of the crystal structure.

Collection of Data

The crystals were fine hexagonal needles, grown by the slow cooling of aqueous solutions of TESM. Since the

absorption coefficient is small for $\text{MoK}\alpha$ radiation, a fairly large crystal was selected for use with the Hilger and Watts Linear Diffractometer (Arndt and Phillips, 1961) in order to obtain fairly large counts. The crystal dimensions were approximately $0.4 \times 0.4 \times 0.8$ mm. Weissenberg photographs of the reciprocal lattice layers were taken about the needle (c) axis using $\text{CuK}\alpha$ radiation. These were taken in order to check the space group absences and to facilitate setting the crystal on the diffractometer. The a axial length was measured on a zero-layer Weissenberg photograph, and both a and c axes were measured from Precession photographs of the $h\ 0\ l$ net. These cell dimensions agreed well with the cell dimensions obtained from the diffractometer.

It was decided to use the hexagonal representation of the space group instead of the rhombohedral representation. Two-thirds of the general hexagonal reflexions are systematically absent ($h\ k\ l$ is only present for $-h + k + l = 3n$; $h\ \bar{h}\ 0\ l$ only for $l = 2n$). Thus to collect the data on the hexagonal reciprocal lattice would have meant that two-thirds of the time would be spent measuring the intensities of systematically-absent reflexions. For this reason another reciprocal cell was chosen with which all the reciprocal lattice points were

systematically-present.

The matrices required to transform the "diffractometer" indices to the corresponding "hexagonal" indices are given below:

$$\begin{bmatrix} h \\ k \\ l \end{bmatrix} = \begin{bmatrix} 2 & 1 & 0 \\ \bar{1} & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \times \begin{bmatrix} h \\ k \\ l \end{bmatrix} + \begin{bmatrix} a \\ 0 \\ 0 \end{bmatrix}$$

"Hexagonal"

"diffractometer"

where $a = 0$ for layers 0, 3, 6, 9 and 12;

$= 1$ for layers 1, 4, 7, 10 and 13;

$= 2$ for layers 2, 5, 8 and 11.

In the Linear Diffractometer there is an analogue of the reciprocal lattice in two machined slides which are positioned by lead screws in order to set the counter to the correct angle for any reflexion in any net up to approximately $2\theta = 60^\circ$. The layer line or net is set by a third slide which is at right angles to the plane of the other two. If the reciprocal axes are oblique, the first two slides can be set at the appropriate (γ^*) angle. If the other two angles (α^* and β^*) are also oblique, this is met by moving the origin of the analogue reciprocal net away from the rotation axis by the appropriate reciprocal distance. The diffractometer provides not only a print-out of the reflexion indices

and all the counts obtained but also a punched paper-tape containing this information which can be processed directly by computer.

Crystal Setting on the Diffractometer

The crystal was mounted and adjusted on a Precession camera, the needle-axis of the crystal being co-axial with the rotation-axis. The crystal and goniometer head were then mounted on the diffractometer, and the crystal adjusted to rotate on the rotation axis. The c^* vertical slide was then raised to the (006) position with both a^* and b^* slides in the zero position. Radiation from the axial reflexions reaches the counter independently of the rotation angle of the crystal. The output from the scintillation counter was fed to the count-rate meter, and the goniometer axes were adjusted slightly until it was seen from the count-rate meter that the count-rate was independent of the rotation angle. (This is similar to adjusting a wheel until it runs true). The vertical slide position was then altered slightly to obtain maximum count-rate. The vertical slide was then returned to the zero position.

From the zero layer Weissenberg photograph a large axial reflexion (500) was chosen. The a^* axis was set to the appropriate reciprocal distance, and the crystal

and goniometer head rotated slowly until the reflexion was located. The size of reciprocal cell unit was then checked, by comparing the calculated position of the ($\bar{3}00$) reflexion with the position of peak intensity. The a^* and b^* slides were then set at 60° , and the positions of axial reflexions on the b^* axis were checked.

All reflexions are measured on the Linear Diffractometer by having the counter stationary, and rotating the crystal about its rotation axis until the reciprocal lattice point cuts the reflecting circle. Subsequent measurements are made by repeating a small angular rotation of the crystal. The angular range of rotation through which the (600) reflexion was observable was about 0.5° . This range increases as the distance of the reflexion from the point of rotation of the reciprocal net decreases. The oscillation range was set at 3.5° .

Whenever the diffractometer was set to collect the next layer, the setting of the rotation axis was checked on the (006) reflexion, and the intensity of the (600) reflexion was monitored to guard against unnoticed decomposition of the crystal in the continual intense X-ray radiation. No deterioration of the TESM crystal was observed throughout the data collection. Very large intensities were re-measured at low power.

Processing of Reflexion Data

The data processing proceeded in four stages:

1. The intensities were corrected for background and for white radiation by the Ross filter technique (1926) by a computer program written by J. G. Sime (1966). This also detected punching errors in the paper tape obtained from the diffractometer. Measurements which failed to pass the program's tests were scrutinised, and compared with the paper print-out from the diffractometer. Those which could be corrected, and the re-scaled intensities which had been measured at low power were then added to the intensities which had already passed the tests.
2. The reflexions of the primitive "diffractometer" reciprocal cell were now re-indexed in terms of the space group's hexagonal cell. This entailed a rotation of the axes and a different translation of the origin for each net of reflexions. Further equivalent reflexions measured in different parts of reciprocal space had different indices. These were re-indexed in terms of their equivalent reflexions in the chosen asymmetric position of reciprocal space. They were then sorted into order, by indices, and the intensities with the same index were averaged. If agreement between the intensities was bad, and there were more than two estimates of the

intensity of the reflexion, the value farthest from the average was eliminated. The remaining intensities were averaged and the agreement of the remaining intensities was checked. The individual intensity measurements were always printed out, as well as the averaged intensity values, so that the processing could be checked by hand. Two Algol computer programs were written to carry out the data processing described above.

3. The intensities were reduced to $|F_o|$ structure amplitudes, a straight-forward operation since the reflexion geometry of the Linear Diffractometer is the same as that of the equi-inclination Weissenberg camera.

3.3. Course of the Structure Analysis

Calculation of Fourier Maps

At the time when the structure was solved, the Fourier Program available could not be used directly to calculate any Fourier map in the rhombohedral or hexagonal space groups. It was decided to treat the summation as though the space group were $P1$ ($P\bar{1}$ for the Patterson map). To do this it is necessary to generate all the structure factors throughout a hemisphere of reciprocal space. This was done by generating all the equivalent structure amplitudes of the Laue Group $\bar{3}m$, and the phases for these amplitudes, required for electron-density

calculations, were obtained by calculating structure-factors for all the generated amplitudes. The generation of the set of equivalent structure-amplitudes was carried out by a computer program written for this purpose.

Structure Solution

The main feature of interest in the solution of this structure was the extent to which it was carried out automatically on the ASS system.

A sharpened three-dimensional Patterson map was calculated, and was searched for peaks by means of the "Fourier Search" program described in Chapter 2. The 3D-Patterson function was calculated because it was desired to check that the equivalent reflexions generated by the special computer program were correct, and to see if there was any possibility of a disordered structure (c.f. TMSM). The map would also indicate whether the index transformations had been carried out correctly. There are two Harker Sections at $w = 0$ and $w = \frac{1}{2}$.

Since R3c is a polar space group the origin of the co-ordinate axis is arbitrary in the z direction, and therefore the sulphur atom z-coordinate was set at zero. The x- and y-coordinates of the heavy-atom were easily obtained from the peak coordinates of the largest non-origin peaks on the Harker sections.

Structure factors calculated for the sulphur atom alone gave an R-factor of 28%. A 3 D electron-density map was then calculated and was searched for peaks by the Fourier Search program. Since the Fourier program could accommodate only rectangular or square sections, considerably more than the asymmetric portion of the Fourier was calculated. The list of peaks obtained from the searching program was examined, and those outside the asymmetric part of the Fourier were discarded. Instead of drawing up the map, the "atom testing" program of the ASS system was used. Although the z-coordinate of the heavy-atom is arbitrary, the map phased by the heavy-atom alone does not have pseudo-symmetry, because the arrangement of the atoms does not exhibit any symmetry above that of the space group $R\bar{3}c$. This made the map relatively simple to interpret, and the program quickly selected 5 peaks out of a list of 7 peaks which were presented to it.

Structure factors calculated for these 5 atoms, plus the sulphur atom, gave an R-factor of 16% over all the reflexions, and the structure has subsequently been refined successfully. No atoms, other than hydrogen atoms, were found on the difference map.

3.4. Least-Squares Refinement

R. Pollard's Least-Squares program (1966) was used for a preliminary refinement. Each reciprocal lattice layer was given an individual scale factor, and each atom a single isotropic temperature parameter. The observations were all given unit weight. After 3 cycles (full matrix) the R-factor was reduced to 10%.

Cruickshank's (1965) Least-Squares program was then used to continue the refinement. Templeton (1965) has shown that if the observations are taken with the crystal rotating about only one axis, and the data from each of the individual reciprocal lattice layers have different scale factors which must be refined, then there is insufficient information to refine all the atoms fully anisotropically. The U_{jj} temperature parameters corresponding to the rotation axis are indeterminate. Since the monitoring of the intensity of one of the strong zero-layer reflexions had disclosed no apparent fall-off of intensity caused by radiation damage, and since the same counting-times had been used for all the data, the different layers were assumed to be on the same relative scale. Only one overall scale factor was employed.

The arbitrary origin in the z-direction was constrained, (Templeton, 1960) by taking the sulphur atom,

z-coordinate as zero. This coordinate was not refined. If the origin is not constrained, the molecule may "walk" through the unit cell. Cruickshank (1965) suggested that the above procedure gives more rapid convergence than the alternative one of refining all the atomic coordinates and altering the sulphur atom z-coordinate back to zero while applying the opposite of this shift to all the other atoms. No convergence difficulties were experienced at any stage of the refinement.

The anisotropic refinement was commenced with unit weights. The R-factor did not drop below 12.6%, however, and the U_{33} temperature parameters all became unreasonably large. It was found on re-examining the scale-factor for the individual layers which had been obtained by the isotropic refinement that there was no smooth trend in the values. Two of the higher layers had irregularly large values.

After considerable thought it was decided not to refine the U_{33} temperature parameters, but to fix these at their isotropic values, and to refine the individual scale-factors. Refinement was then commenced. The weighting scheme employed was:-

$$W = K \exp - (A + Bf_0 + Cf_0^2 + D(\chi^2/\sin^2\theta) + E(\chi^4/\sin^4\theta) + F(\sin^2\theta/\lambda^2))$$

where K is an arbitrary scale-factor, and the constants A , B , C , D , E and F are fitted by Least-Squares by the program described in Chapter 2. This produced a weighting scheme in which different batches of $\overline{W \Delta^2}$ by $|F_o|$ and by $\sin^2\theta/\lambda^2$ did not differ by more than $\pm 10\%$. Three consecutive cycles of refinement were then carried out. The estimated standard deviations of the parameters improved greatly, and the two chemically-equivalent sulphur-oxygen bonds which had differed by $\sim 0.1\text{\AA}$ now differed by only $\sim 0.01\text{\AA}$.

3.5. Difference Map and Location of Hydrogen Atoms

Anisotropic structure-factors were calculated using R. Pollard's structure-factor program, and a difference-map was calculated. All the hydrogen atoms were located without difficulty; this also implies that the thermal vibration parameters could not be large in the direction of the c -axis. No large positive or negative peaks were observed, although the fluctuating background was usually in the region of ± 0.2 electrons / \AA^3 .

Positions for the two methylene hydrogen atoms were then calculated using G. Ferguson's hydrogen-placing program. The calculated coordinates for these atoms were used in the subsequent refinement. The coordinates of the hydrogen atom of the central carbon atom and those

of the methyl group were those obtained from the difference map.

3.6. Polar coordinate shift caused by uncorrected Anomalous Dispersion

It has been pointed out (Templeton, 1966; Cruickshank and MacDonald, 1966) that the neglect of the anomalous dispersion correction in structure-factor-least-squares refinement can lead to a serious error in the polar coordinates of the anomalously scattering atoms. This occurs in cases where the data have been collected from only one reciprocal hemisphere. Cruickshank has estimated that the correction in the direction of the polar axis would be of the order of 0.01\AA for a sulphur atom scattering $\text{MoK}\alpha$ radiation. This is approximately three to four times the estimated standard deviation of the sulphur atom in TESM, so that the anomalous dispersion correction was required. The data had been collected only from the upper half of the reciprocal lattice, and the equivalent intensities had been averaged. All these equivalent reflexions are, however, of the same "hand" with regard to the anomalous dispersion correction.

In the space group $R\bar{3}c$ either the molecule or its enantiomer are present although the TESM chemical molecule is not optically active, and the activity cannot be

resolved since one enantiomer can easily be changed into the other by rotation about the sulphur-central carbon single bonds.

In Fig. 5a, F , \bar{F} are structure factors which would ordinarily be related by Friedel's law.

F'_A , \bar{F}'_A are the real components of the scattering vector caused by the anomalously scattering atom.

F''_A , \bar{F}''_A are the imaginary components of the scattering vector caused by the anomalously scattering atom.

F_R , \bar{F}_R are the real scattering components of the other ordinary atoms.

Fig. 5b shows the modulus of \bar{F} drawn in the first quadrant for comparison with F .

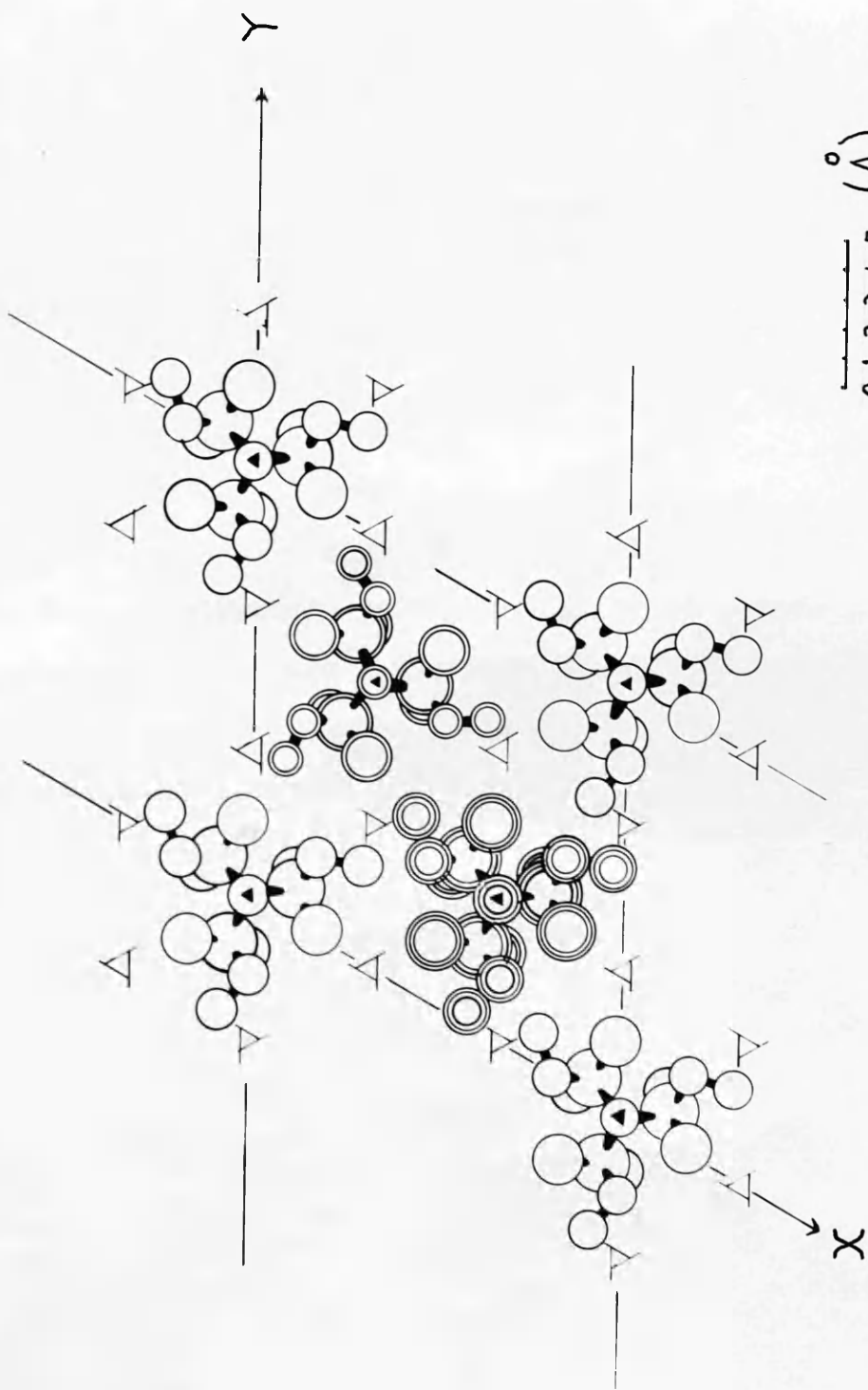
In structures where only one reflexively dissymmetric enantiomer is present, and data from only one half of reciprocal space has been collected, models of both enantiomers will give equally good agreement in the absence of anomalous dispersion, and with either of these the data collected may have been either F_{hkl} or $F_{\bar{h}\bar{k}\bar{l}}$.

Thus four possibilities arise:-

FIG. 3

Projection of structure down z-axis onto xy plane.

The molecules shown are sited between $z = 0.5$ and $z = 0$. The molecules with the largest z-coordinates are shown with three circles; those with the smallest z-coordinates are given only one circle.



0 1 2 3 4 5 (Å)

FIG. 4

Projection of structure perpendicular to yz plane.
The molecules shown are all sited between $x = 0.5$
and $x = 1$. Those molecules nearer $x = 1.0$ are
drawn as double circles.

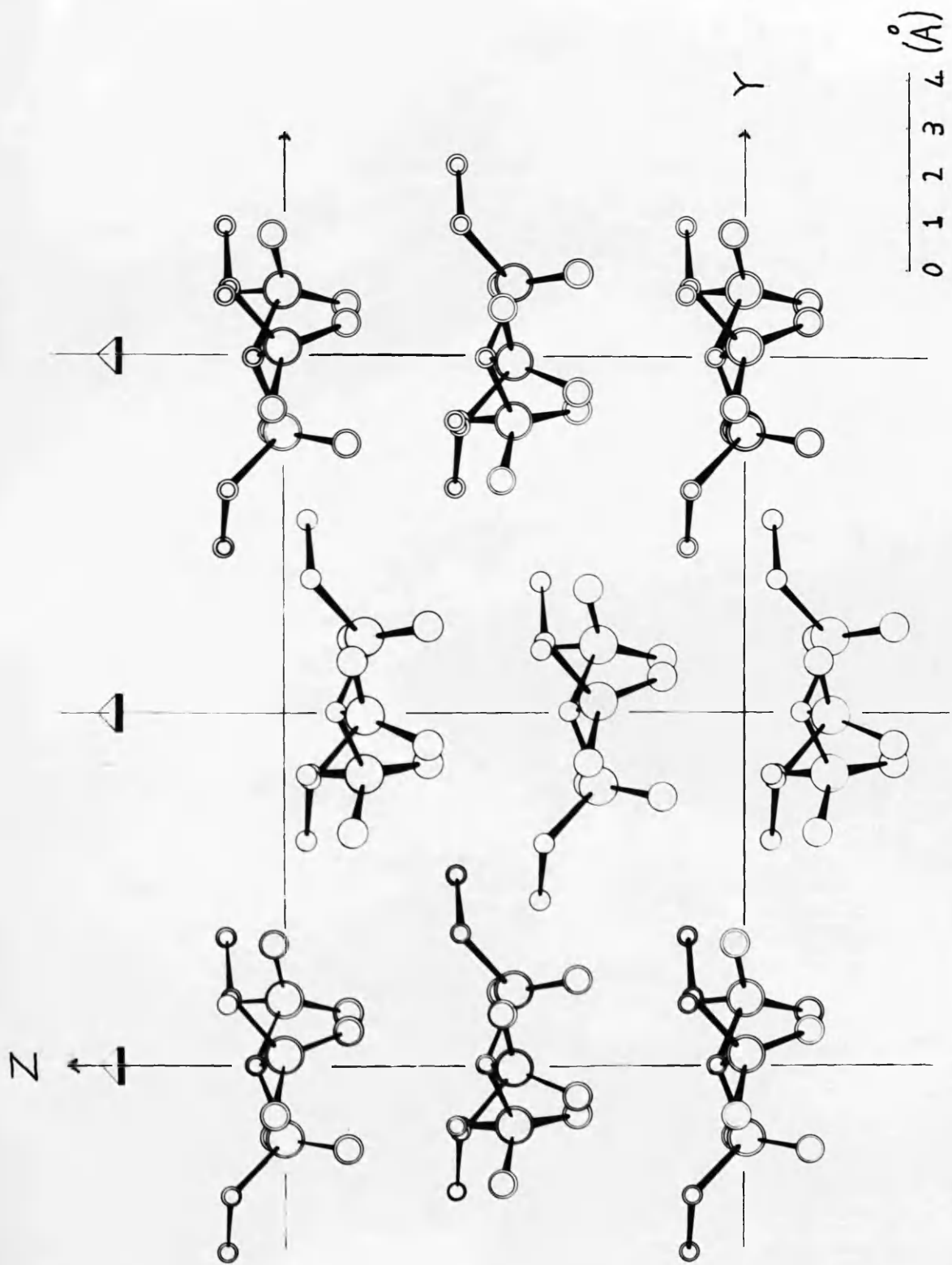


FIG. 5

The effect of anomalous dispersion on the observed structure factors.

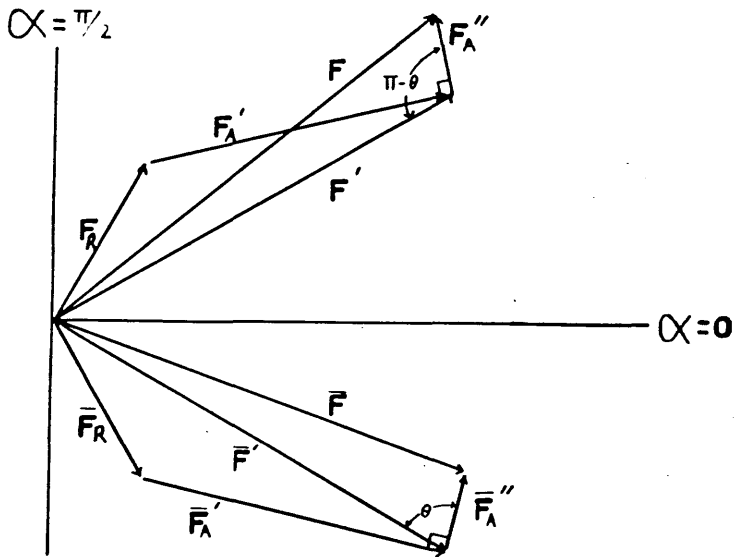


Fig 5a

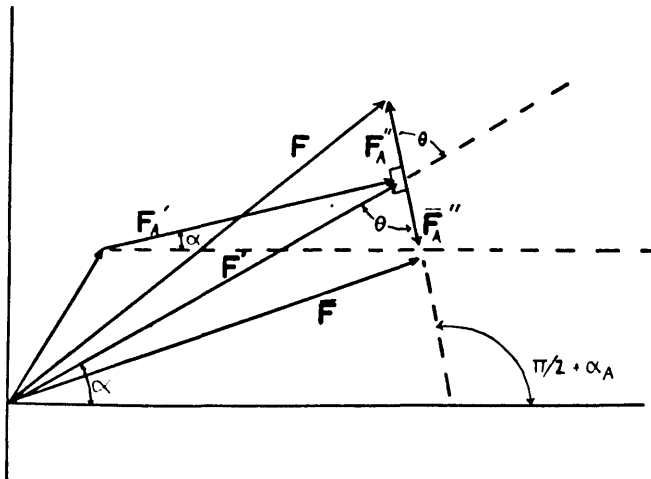
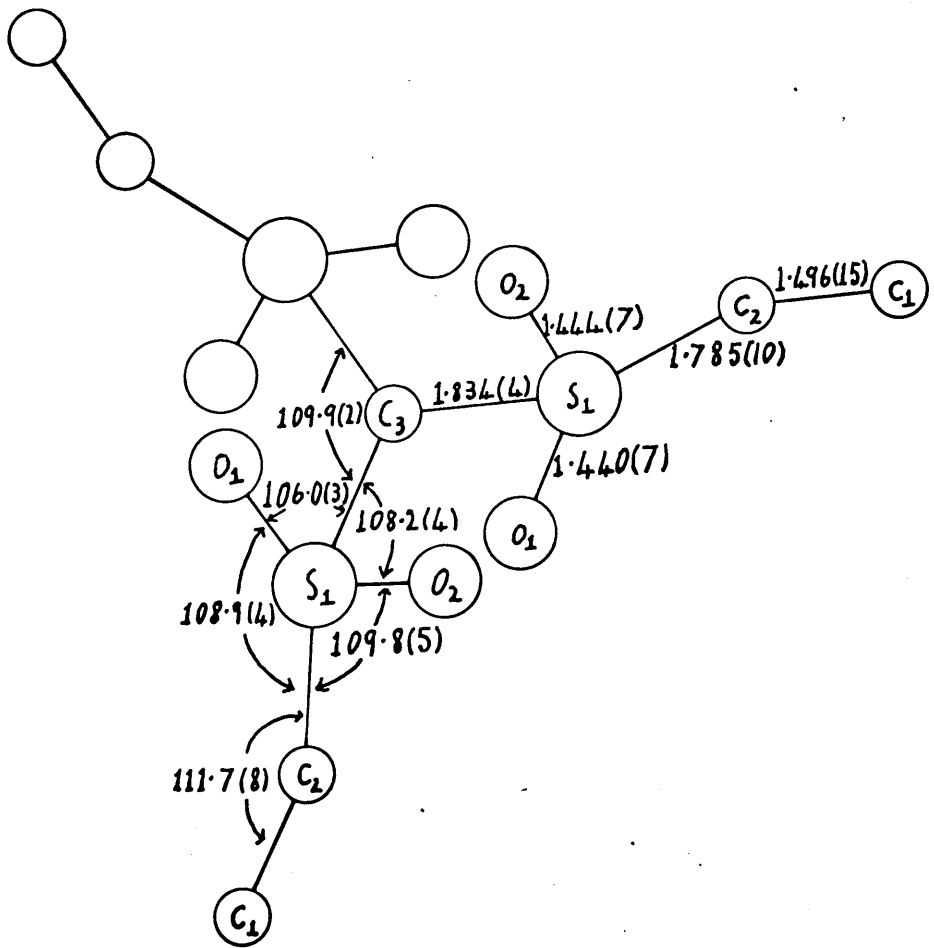


Fig 5b

FIG. 6

Important bond-lengths and bond-angles



<u>I</u>	<u>II</u>
Molecule	Molecule
Fhkl	F \bar{h} k \bar{l}
<u>III</u>	<u>IV</u>
Enantiomer	Enantiomer
Fhkl	F \bar{h} k \bar{l}

When anomalous dispersion occurs, a choice can be made as to whether the data are |Fhkl| or |F \bar{h} k \bar{l} |, but the absolute configuration cannot be determined unless data are collected from both halves of reciprocal space.

In this case both enantiomers are present whenever the substance is dissolved by any solvent, so only the two possible ways of indexing the data have to be considered, because the absolute configuration of the molecule is of little interest.

The comparison between these was made by doing two separate refinements. The change in indexing to F \bar{h} k \bar{l} was simulated by changing the sign of f'' in the structure-factor least squares calculation, where f'' is the imaginary component of the anomalous atom's scattering-factor.

The results of the refinements are shown below:-

R%	8.52	8.72
S-(0(1))	1.440 ⁺⁷	1.433 ⁺⁷
S-0(2)	1.444 ⁺⁸	1.456 ⁺⁸

The lengths of the two sulphur-oxygen bonds are given because they are sensitive to the polar shift correction, S-O(1) being almost perpendicular to the c-axis, S-O(2) being parallel to the c-axis. The polar shift correction will thus have maximum effect on S-O(2) length and much less effect on the S-O(1) distance.

The final coordinates and parameters given in this chapter are those which correspond to the data being indexed as |Fhkl|.

3.7. Description and Discussion of Structure of TESM

The molecular and crystal structure of TESM are shown in Fig. 3 and Fig. 4; the more important interatomic distances and angles are shown in Fig. 6 and Tables 8 and 9. For comparison, the interatomic distances of TMSM (Silverton, 1966) are shown in Table 10. The interatomic distances and angles were calculated by K. W. Muir's (1965) program.

The configuration of the bonds of the central carbon atom is clearly tetrahedral. This is similar to the configuration of TMSM, and differs from the situation in the ammonium salt of TMSM which is planar (Hoogsteen, 1957). The two chemically-equivalent sulphur-oxygen bonds are:

S-O(1) 1.440 \pm 7

S-O(2) 1.444 \pm 8

These are virtually identical, and compare very well with average sulphur-oxygen distance of 1.435 Å in TMSM, and in Dichlorodiphenyl sulphone (DCDPS) (Sime and Abrahams, 1960). In both TESM and TMSM the two sulphur-oxygen bonds point in different directions, one bond-direction being almost parallel to the c-axis, the other being almost perpendicular to the c-axis. The anomalous dispersion correction affects the length of the bond parallel to the polar-axis much more than the other bond-length. The correction was applied to TESM but not to TMSM.

The O - S - O bond angle is 119.1° which compares with 120.4 \pm .4° in DCDPS, 119.3 \pm .9° in TMSM, and 118.4° in N-methyl-2:2-dimethyl sulphonyl vinylidene amine (NMDSVA) (Wheatley, 1954).

There are two distinct carbon-sulphur bond lengths in TESM, sulphur-C central and sulphur-C ethyl.

Sulphur-C central

This bond length is 1.834 \pm 4Å. It compares with 1.83 \pm 1Å in TMSM, 1.726Å in NMDSVA and 1.70 Å in NH₄⁺ TMSM⁻. Partial double-bond character of the sulphur-carbon bond is entirely consistent with the chemical

formulae of these last two compounds.

Sulphur-C ethyl

This bond length is $1.785 \pm 10^{\circ}$ and compares with the corresponding sulphur-methyl bond in TMSM with a length of $1.73 \pm 2^{\circ}$, and with 1.71° in NMDSVA. Corrections for vibration have not been applied to any of these bond-lengths, and this correction would lengthen all the bonds slightly. It may be noted that the C-C bond of TESM is slightly shorter than the accepted C-C single bond-distance. Nevertheless, there can be no suggestion of double-bond character in these C-S bonds in TESM and TMSM although both are shorter than the standard C-S single bond distance of 1.815° proposed by Abrahams (1956). One substance having the standard S-C single bond distance is DCDPS where the group attached to the carbon atom is powerfully electron-withdrawing. This may well lengthen the bond.

Molecular Packing

The packing of the molecules is shown in Fig. 3 and Fig. 4. As in TMSM the molecules are packed in "stacks" or columns, the axis of each stack being one of the 3-fold axes of the space group. As mentioned earlier, in the space group $R\bar{3}c$ the columns consist of the molecule, then the molecule rotated to its other orientation, alternately

along the length of the stack. Although the molecules in the crystal are reflexively dissymmetric, with TESM and TMSM molecules the enantiomer is produced by a rotation of the R-SO₂ group about the S-C central bonds, and there can be little, if any, energy barrier between the two enantiomers. Thus the enantiomers could never be resolved chemically. The molecules of the column are not bound together significantly by hydrogen bonding along the 3-fold axis. The distance between the central carbon atom and the three "axial" oxygen atoms of the next molecule in the stack (C.....O(2)) is 3.2Å.

Three stacks are grouped together about 3₁ or 3₂ screw axes. In Fig. 3, the projection of the lower half of the unit cell down the z-axis on to the xy plane, it can be seen that the molecules in different stacks intermesh like three-toothed gear wheels. Fig. 4 shows the projection normal to the yz plane. Here the molecules of the different columns interleave with each other.

3.8. TESM and TMSM

It was anticipated that the structures of TESM and TMSM would be isomorphous, or that TESM would be isostructural with either the "major" or "minor" TMSM structures. In spite of the similarity of the actual

coordinate values it can be seen from the Appendix that the structure of TESM differs from both TMSM structures. It can be seen that the relationship between the structures is that the TESM molecule is inverted with respect to the TMSM molecule. This corresponds to an interchange of a and b-axes, or to the replacement of c by \bar{c} -axis. No space group transformation can be applied to TESM coordinates to convert them into approximations to the TMSM coordinates.

Further, when the a^* -axis has been chosen in the reciprocal lattice, the general absences and the symmetry of the Laue Group $\bar{3}m$ define the directions of the b^* -and c^* -axes. It is not possible to interchange b^* and c^* without changing c^* to \bar{c}^* . The Patterson maps of both TESM and TMSM were calculated during the course of this investigation. Their features are different, and the positions of the sulphur atoms deduced from them are different. The electron-density maps phased from the sulphur-atom-structure-factors give essentially the known structures, the atomic coordinates of which are shown in Tables 12 and 13. Indeed, since the chemical molecules both have the same tetrahedral shape and the same 3-fold axis, and the molecular centres must each occupy a special position, the "similarity" of the coordinates is not surprising.

TABLE 3.1

FRACTIONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
S(1)	0.1329 ± 1	0.0796 ± 2	0.0000 ± 0
O(1)	0.1832 ± 5	0.0206 ± 5	0.0253 ± 7
O(2)	0.1292 ± 5	0.1119 ± 6	-0.1375 ± 7
C(1)	0.3012 ± 9	0.2382 ± 12	0.1206 ± 13
C(2)	0.1864 ± 7	0.1888 ± 7	0.1109 ± 10
C(3)	0.0000 ± 0	0.0000 ± 0	0.0610 ± 10
H(1)	0.0000 ± 0	0.0000 ± 0	0.187 ± 21
H(2)	0.175 ± 16	0.248 ± 17	0.065 ± 21
H(3)	0.185 ± 20	0.195 ± 22	0.186 ± 30
H(4)	0.304 ± 22	0.244 ± 20	0.074 ± 30
H(5)	0.307 ± 20	0.296 ± 20	0.202 ± 33
H(6)	0.316 ± 20	0.175 ± 19	0.173 ± 23

TABLE 3.2

FRACTIONAL COORDINATES AND F.S.D.

ATOM	X	Y	Z
S(1)	0.1329 ± 1	0.0796 ± 2	0.0000 ± 0
O(1)	0.1832 ± 5	0.0206 ± 5	0.0253 ± 7
O(2)	0.1292 ± 5	0.1119 ± 6	-0.1375 ± 7
C(1)	0.3012 ± 9	0.2382 ± 12	0.1206 ± 13
C(2)	0.1864 ± 7	0.1888 ± 7	0.1109 ± 10
C(3)	0.0000 ± 0	0.0000 ± 0	0.0610 ± 10

TABLE 3.3

COORDINATES AND E.S.D. IN ANGSTROMS
WITH RESPECT TO THE CRYSTAL AXES

ATOM	X	Y	Z
S(1)	1.9887 ± 21	1.1914 ± 30	0.0000 ± 0
O(1)	2.7416 ± 73	0.3084 ± 81	0.2483 ± 68
O(2)	1.9327 ± 81	1.6743 ± 87	-1.3492 ± 69
C(1)	4.5079 ± 142	3.5644 ± 175	1.1834 ± 127
C(2)	2.7896 ± 103	2.8251 ± 109	1.0881 ± 94
C(3)	0.0000 ± 0	0.0000 ± 0	0.5984 ± 99

ORTHOGONAL COORDINATES

The orthogonal axes (X, Y, Z) are defined as follows:

X is parallel to a^*

Y is parallel to b

Z is parallel to c

TABLE 3.4

ORTHOGONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
S(1)	1.7223 \pm 18	0.1970 \pm 28	0.0000 \pm 0
O(1)	2.3743 \pm 64	-1.0624 \pm 70	0.2483 \pm 68
O(2)	1.6738 \pm 70	0.7079 \pm 73	-1.3492 \pm 69
C(1)	3.9040 \pm 123	1.3104 \pm 167	1.1834 \pm 127
C(2)	2.4159 \pm 89	1.4303 \pm 106	1.0881 \pm 94
C(3)	0.0000 \pm 0	0.0000 \pm 0	0.5984 \pm 99

TABLE 3.5

FRACTIONAL COORDINATES AND TEMPERATURE FACTORS
FOR HYDROGEN ATOMS

ATOM	X	Y	Z	U
H(1)	0.000	0.000	0.187	0.0152
H(2)	0.175	0.248	0.065	0.0487
H(3)	0.185	0.195	0.186	0.0209
H(4)	0.304	0.244	0.074	0.0150
H(5)	0.306	0.296	0.202	0.0799
H(6)	0.316	0.174	0.173	0.0715

TABLE 3.6

ANISOTROPIC TEMPERATURE FACTORS AND E.S.D.

ATOM	U11	U22	U33	2U23	2U31	2U12
S(1)	0.0302 9	0.0339 10	0.0215 0	0.0041 15	0.0090 11	0.0288 14
O(1)	0.0426 29	0.0569 37	0.0333 0	0.0086 48	0.0078 44	0.0570 55
O(2)	0.0545 40	0.0434 36	0.0271 0	0.0176 52	0.0218 46	0.0426 72
C(1)	0.0493 54	0.0806 89	0.0431 0	-0.0169 102	-0.0208 83	0.0348 114
C(2)	0.0384 41	0.0352 43	0.0345 0	-0.0028 61	-0.0059 60	0.0183 66
C(3)	0.0261 28	0.0261 8	0.0174 0	0.0000 0	0.0000 0	0.0261 28

TABLE 3.7

PRINCIPAL VALUES AND DIRECTIONS OF
VIBRATION TENSORS
WITH RESPECT TO ORTHOGONAL AXES

ATOM	^{o2} U A	D1	D2	D3
S(1)	0.0196	-0.3909	-0.0063	0.9204
	0.0362	0.1848	-0.9802	0.0718
	0.0319	0.9017	0.1981	0.3843
O(1)	0.0577	0.5100	0.8421	0.1754
	0.0319	-0.3560	0.0210	0.9342
	0.0384	0.7830	-0.5389	0.3105
O(2)	0.0219	-0.3518	-0.2302	0.9073
	0.0606	0.8850	-0.3976	0.2423
	0.0468	0.3049	0.8882	0.3436
C(1)	0.1021	0.1495	-0.9881	0.0365
	0.0344	0.6122	0.1215	0.7813
	0.0566	-0.7764	-0.0944	0.6231
C(2)	0.0556	-0.5675	0.8192	0.0828
	0.0294	0.7622	0.4846	0.4292
	0.0355	0.3115	0.3067	-0.8994
C(3)	0.0261	1.0000	0.0000	0.0000
	0.0261	0.0000	1.0000	0.0000
	0.0174	0.0000	0.0000	1.0000

TABLE 3.8

BOND-LENGTHS WITH E.S.D.

S(1) - O(1)	1.440	\pm	7
S(1) - O(2)	1.444	\pm	7
S(1) - C(2)	1.785	\pm	10
S(1) - C(3)	1.834	\pm	4
C(2) - C(1)	1.496	\pm	15
C(3) - H(1)	1.24	\pm	21

TABLE 3.9

BOND-ANGLES WITH E.S.D.

O(1) - S(1) - O(2)	119.1 \pm 4
O(1) - S(1) - C(2)	108.9 \pm 4
O(1) - S(1) - C(3)	106.0 \pm 3
C(1) - C(2) - S(1)	111.7 \pm 8
C(2) - S(1) - C(3)	103.9 \pm 4
S(1) - C(3) - S(1)	109.9 \pm 2
S(1) - C(3) - H(1)	109.0 \pm 2

TABLE 3.10

BOND-LENGTHS WITH E.S.D.

(TMSM)

S(1) - O(1)	1.427 \pm 17
S(1) - O(2)	1.442 \pm 12
S(1) - C(1)	1.729 \pm 17
S(1) - C(2)	1.830 \pm 10

TABLE 3.11

Final observed and calculated structure factors.

TABLE 3.12

TMSM major Structure, with S at $z = 0.0000$

	x	y	z
S	0.1525	0.1040	0.0000
O(1)	0.2217	0.0469	-0.0219
O(2)	0.1415	0.1401	0.1399
C(1)	0.2061	0.2230	-0.1159
C(2)	0.0000	0.0000	-0.0597

TABLE 3.13

TMSM minor Structure with S at $z = 0.0000$

	x	y	z
S	0.1040	0.1525	0.0000
O(1)	0.0469	0.2217	0.0219
O(2)	0.1401	0.1415	-0.1399
C(1)	0.2230	0.2061	0.1159
C(2)	0.0000	0.0000	0.0597

APPENDIX

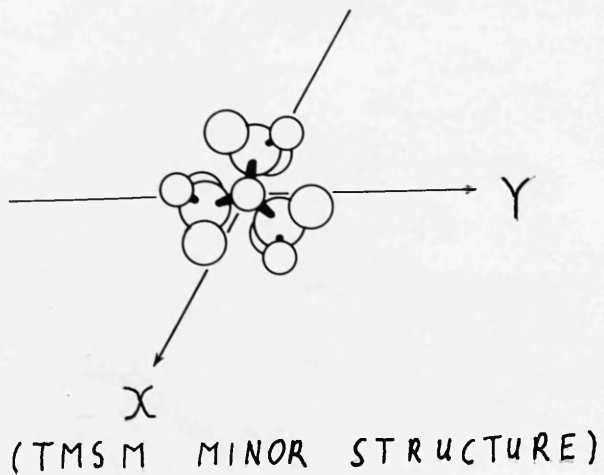
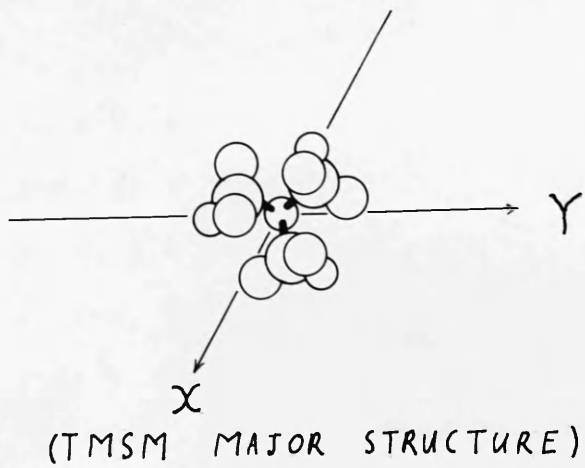
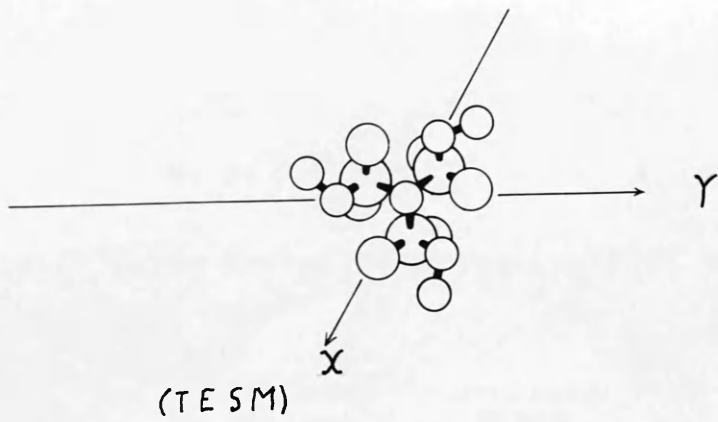
COMPARISON OF TESM with TMSM (major) and TMSM (minor) STRUCTURES

The fractional coordinates of TESM, TMSM (major) and TMSM (minor) are given in Tables 1, 12 and 13 respectively. These tables give the molecular coordinates adjusted so that the (arbitrary) z-coordinate of the sulphur atom is zero in all three cases. Tables 12 and 13 therefore do not show the correct z-values of TMSM (major) relative to TMSM (minor), but the purpose here is to compare both these structures independently with that of TESM. In making this comparison, any of the equivalent positions of the molecules in the unit cell may be taken, the origin of the coordinate system may be moved in the z direction, and, since the absolute configuration of the molecules is unknown, a centre of symmetry may be applied, in the attempt to convert the TESM coordinates into others analogous with one or both of the TMSM coordinates.

Let the TESM fractional coordinates be x , y , z .
The following approximate relationships then occur:

FIG. 7

Comparison of TESM structures with TMSM (Major)
and TMSM (Minor) structures.



TESM	TMSM (major)
x, y, z	x, y, \bar{z}
TESM	TMSM (minor)
x, y, z	y, x, z

Listed below are all the "equivalent" TMSM coordinates:-

Actual Coordinates	Space Group Equivalent	z-origin shift	hkl to $\bar{\bar{\bar{hkl}}}$
x, y, z	x, y, z		
	\bar{y} , x-y, z		
	y-x, \bar{x} , z		
	\bar{y} , \bar{x} , $\frac{1}{2} + z$	\bar{y} , \bar{x} , z	y, x, \bar{z}
	x, x-y, $\frac{1}{2} + z$		
	y-x, y, $\frac{1}{2} + z$		

CHAPTER 4

THE STRUCTURE DETERMINATION OF POTASSIUM HYDROGEN
DIANISATE

Contents

- 4.1. Introduction
- 4.2. Experimental
- 4.3. Course of the Analysis
- 4.4. Least-Squares Refinement
- 4.5. Discussion of Structure

THE STRUCTURE DETERMINATION

OF

POTASSIUM HYDROGEN DIANISATE

4.1. Introduction

Acid salts of carboxyl acids are very common substances. The potassium and rubidium acid salts of anisic acid are readily prepared by mixing alcohol-water solutions of the acid with potassium hydroxide or rubidium carbonate, in the ratio of one equivalent of acid to half an equivalent of the metal hydroxide or carbonate. The crystals are produced when the solutions are allowed to evaporate slowly.

Crystal structures of acid salts have frequently been studied by X-ray diffraction techniques (Mills and Speakman, 1961; Bryan, Mills and Speakman, 1963A; Mills and Speakman, 1963B; Golič and Speakman, 1965A; Golič and Speakman, 1965B). They have also been examined by neutron diffraction (Currie, Curry and Speakman, 1967) and their infra-red spectra have been classified by Blinc, Hadži and Novak (1960), Shrivastava and Speakman (1961). The chemical interest is centred on the short hydrogen bonds which occur in some of these compounds.

The infra-red spectra of these two acid salts of anisic acid show the anomaly characteristic of "Type A" acid salts (Shrivastava and Speakman, 1961). That the structures are, in fact, "Type A", has been confirmed by the solution of both structures. In the case of the potassium salt the length of the O O distance in the effectively symmetrical hydrogen bond has been determined with considerable accuracy.

4.2. Experimental

Crystal Data

The following values were established by photographic methods with $\text{CuK}\alpha$ - radiation ($\lambda = 1.5418 \text{ \AA}$).

Potassium hydrogen dianisate, $\text{KH}(\text{C}_8\text{H}_7\text{O}_3)_2$; orthorhombic; $M = 342.4$; $a = 35.77$; $b = 7.05$; $c = 6.16 \text{ \AA}$; $V = 1553 \text{ \AA}^3$; $D_m = 1.43$ (by flotation in bromoform-carbon tetrachloride); $Z = 4$; $D_c = 1.46$; $F_{000} = 712$; space group Pbcn ; absorption coefficient for $\text{Cu K}\alpha$ radiation = 33.2 cm^{-1} ; crystal shape, lath-like crystals (cut for intensity determination).

Rubidium hydrogen dianisate $\text{RbH}(\text{C}_8\text{H}_7\text{O}_3)_2$; $M = 388.8$; Orthorhombic; $a = 35.3 \pm 0.2 \text{ \AA}$; $b = 7.00 \pm 0.10 \text{ \AA}$; $c = 6.40 \pm 0.10 \text{ \AA}$; $V = 1581 \text{ \AA}^3$; $D_m = 1.63 \pm .01$ (by flotation in $\text{CHBr}_3 - \text{CCl}_4$); $Z = 4$; $D_c = 1.62$; $F_{000} = 784$; space group Pbcn ; absorption coefficient $\text{Cu K}\alpha = 47.6 \text{ cm}^{-1}$.

4.3. Course of the Analysis

From the similarity of space group, cell dimensions and the number of formula units per unit cell, it was inferred that the potassium compound was isomorphous with the rubidium one. The space group implies that the metal ion and the acidic hydrogen atom occupy 4-fold positions while the two anisate residues of the chemical formula occupy 8-fold symmetry positions and are therefore

identical, symmetry-related molecules. This situation is typical of "Type A" acid salts.

The structure of rubidium hydrogen dianisate had been solved in two projections (Forgie, 1964) (Skinner, 1950). Since the two crystal structures were believed to be isomorphous, it was decided to work only on the potassium compound which has a much smaller absorption coefficient. Intensity data were measured visually from multiple film Weissenberg photographs, which were taken about both b and c axes, although the crystal shape caused the b-axis photographs to be subject to much greater absorption errors.

The potassium ion was assumed to occupy a 4-fold position of the type $0, y, \frac{1}{4}$. This was checked by calculating a Harker section of the 3D Patterson at $w = \frac{1}{2}$, from which the y-coordinate of the potassium ion was deduced.

Structure factors calculated for the potassium ion alone gave a R-factor of 42%. The position of the heavy-atom is such that the three-dimensional electron-density map calculated using the heavy-atom phased structure amplitude has mirror planes at $z = \frac{1}{4}$ and $z = \frac{3}{4}$. This pseudo-symmetry was destroyed by choosing one of the images of the carboxyl group of the acid residue. A

set of structure factors and an electron-density map were then calculated, and this time it was possible to locate the benzene ring and the methoxyl group.

A third structure-factor-electron-density map calculation was carried out, and coordinates obtained from this map gave an R-factor of 24%.

4.4. Least-Squares Refinement

The initial Least Squares Refinement was carried out with an early version of the Cruickshank and Smith Least Squares program. This early version could not refine batch scale parameters for the different reciprocal lattice layers. Thus layer scaling was done by calculating:

$$k = \frac{\sum_{\text{layer}} |F_c|}{\sum_{\text{layer}} |F_o|} ,$$

with a scale and agreement-factor program written by the author. In the first three cycles unit weights and individual atomic isotropic temperature parameters were employed. The weighting scheme was adjusted to make the value of $\langle \overline{w \Delta^2} \rangle$ (average) constant for batches of observations within different ranges by $|F_o|$. This adjustment was made by altering the parameters p_1, p_2, p_3 in the expression for \sqrt{w} :-

$$\sqrt{w} = \frac{1}{(p_1 + F_o + p_2 F_o^2 + p_3 F_o^3)^{1/2}}$$

until, by trial and error, an approximately "flat" distribution of $\langle \overline{w \Delta^2} \rangle$ was obtained.

"Unobserved" reflexions were given an intensity equal to half the minimum observed intensity, i.e.

$$|F_o|_{\text{unobserved}} = \frac{1}{2} |F_o|_{\text{local minimum}}.$$

This initial refinement was somewhat unsatisfactory; the estimated standard deviations were insufficient to account for the anomalous bond-lengths of the benzene ring (C(2) - C(3) 1.32 Å, C(2) - C(7) 1.44 Å); the scale parameter values had been obtained by a somewhat rough-and-ready method, and the distribution of $\langle \overline{w\Delta^2} \rangle$ by $\sin^2\theta/\lambda$ was not particularly "flat".

A difference map was calculated, from which it was possible to locate the four hydrogen atom peaks of the benzene ring, but not those of the methoxyl group. (Subsequent analysis has shown that this group has high thermal vibration parameters). When layer-scale-parameter refinement became possible with a later version of the Least-Squares Program, it was decided to continue the refinement, and to apply a different weighting scheme. The weighting scheme applied was of the form:

$$w = K \exp - \left\{ A + B\theta + C\theta^2 + D\theta^3 + E \left(\frac{\lambda^2}{\sin^2\theta} \right) + F \left(\frac{\lambda^4}{\sin^4\theta} \right) + G \left(\frac{\sin^2\theta}{\lambda^2} \right) \right\},$$

where K is an arbitrary parameter for scaling the relative weights and the coefficients A, B, C,..... G are fitted by the Least-Squares method by the program described in Chapter 2.

It was possible to obtain fully anisotropic atomic vibration parameters for this structure because data

had been collected about both b and c axes, although, as mentioned earlier, the b axis data were subject to severe absorption errors because of the crystal shape. This refinement produced marked changes in the layer-scale parameters. There was a considerable improvement in the coordinate estimated standard deviations, as would have been expected from the more comprehensive weighting-scheme, but more important than this was the change in the bond-lengths of the benzene ring. Final values are given in Table 8.

A final difference map was calculated, but the missing hydrogen atoms of the methoxyl group could still not be located. There was a complex region of positive and negative density surrounding the heavy-atom position. The difference map, however, displays the transformation in real space of $(F_o - F_c)$ whereas the least-squares procedure minimises a function of $W(F_o - F_c)^2$. The complex region is thus probably caused by the uncorrected absorption and extinction errors of the large low-angle reflexions, which are down-weighted by the weighting scheme. The final R-factor is 15.1%. This may seem high, but it is considered to be reasonable in view of the absorption errors of the $h0l$ and $h1l$ nets, and the large number of "unobserved" reflexions.

4.5. Discussion of Structure

The atomic parameters of the structural model are given in Tables 1 to 5; the more important bond-lengths and angles are given in Tables 6 to 10, along with the more important interatomic contacts. In addition, the molecular bond-lengths and angles are given in Fig. 1. The general packing of the molecules is shown in Fig. 2 and Fig. 3.

This structure is a "Type A" acid salt, and the two acid residues of the gross formula $\text{KH}(\text{An})_2$ are crystallographically equivalent. They cannot be distinguished as H-An (acid) and An^- (anion) and this is the structural criterion used to allocate an acid salt to one of the two classes. The mean of the bond-lengths of the bonds in the benzene ring is $1.393 \text{ \AA} \pm 12^\circ$ and none of the individual bond-lengths differs significantly from this value. The C(1) - C(2) bond-length is somewhat shorter than the normal ($\text{sp}^3 - \text{sp}^3$) C - C single-bond length 1.54 \AA , since the carboxyl carbon atom is sp^2 hybridised. The length agrees well with values found in other similar compounds (Table 12. The average length is 1.495 \AA which is in good agreement with the value found here.

Neither the oxygen of the methoxyl group nor the

FIG. 1

Important bond-lengths and bond-angles.

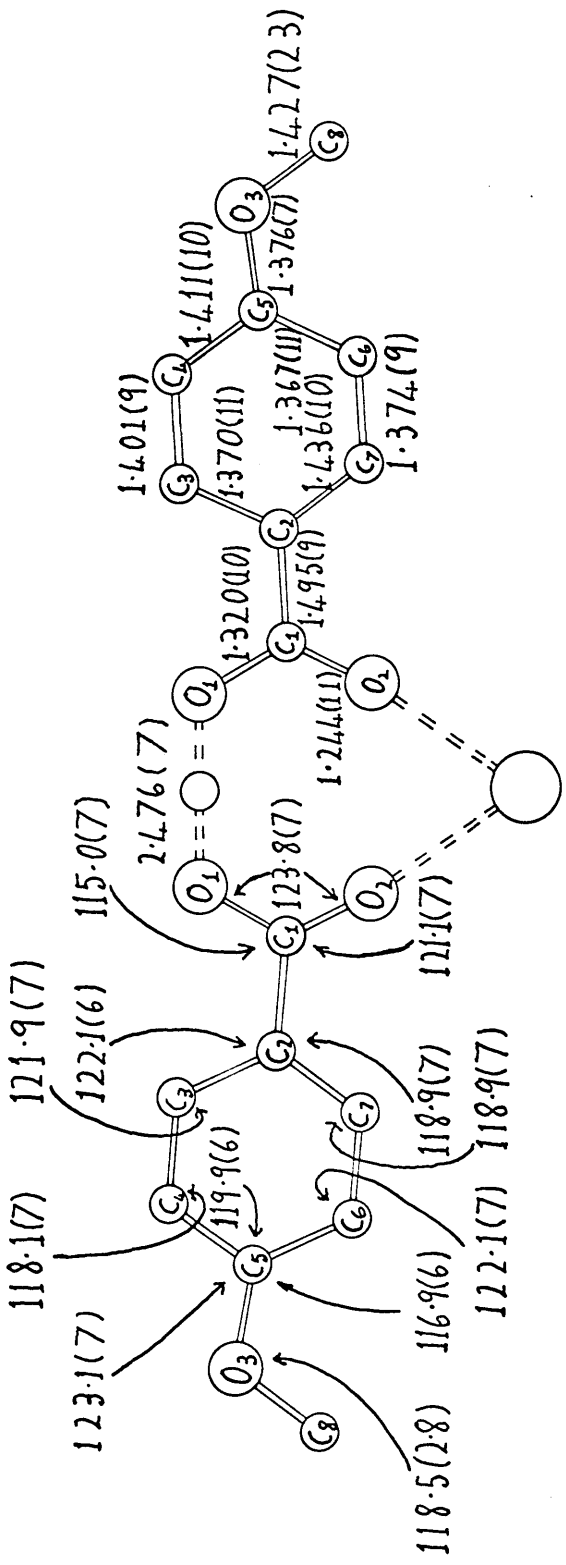


FIG. 2

Projection down the z-axis onto the xy plane of unit cell contents from $z = 0$ to $z = 1$. The atoms of the molecules in the upper half of the unit cell are drawn with double lines.

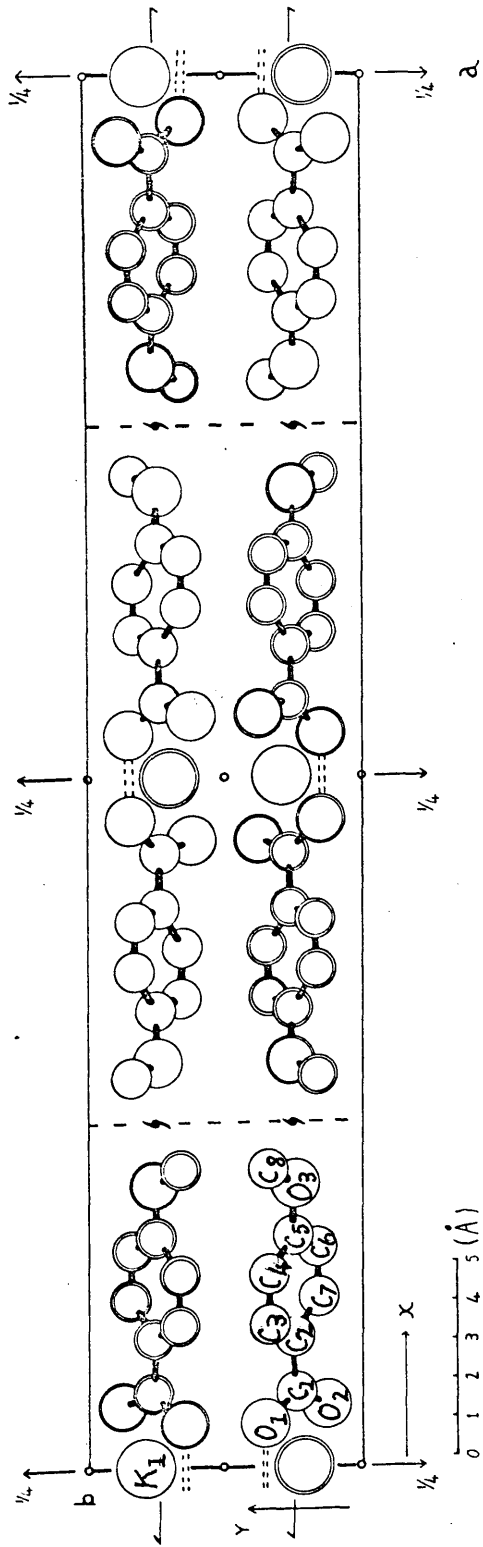
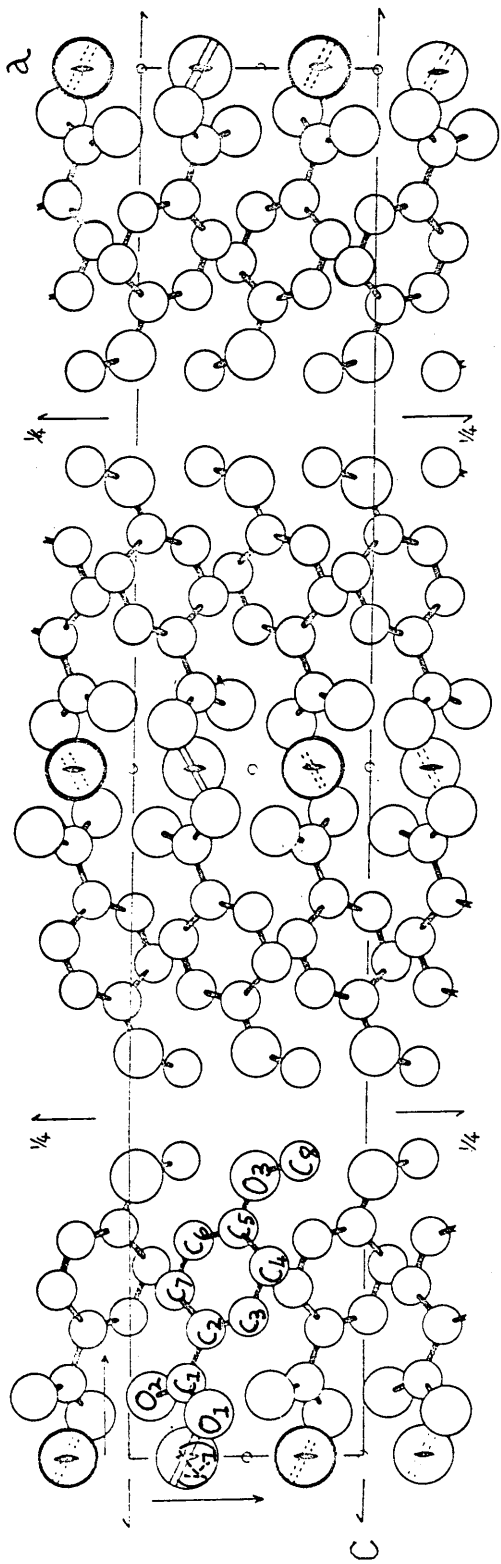


FIG. 3

Projection down the y -axis onto the xz plane of the
unit cell contents from $y = 0$ to $y = 1$.



0 1 2 3 4 5 (Å)

carbon atom of the carboxyl group deviates significantly from the plane of the benzene ring which is given by the parameters:-

$$0.1516X + 0.8592Y - 0.4886Z = 1.063 \quad (1)$$

where X, Y, Z are given in Å with respect to the crystal axes.

The mean plane through the carboxyl group and C2 is given by:

$$0.2102X + 0.6987Y - 0.6838Z = .5361 \quad (2)$$

The angle between planes (1) and (2) is 14.9° . Since the planar configuration would be expected to be stabilised by delocalisation of the anion's extra electron with the benzene ring, this large angular displacement is, at first, surprising. There are no larger substituents in the ortho position of the benzene ring, and the p-methoxy group cannot have any effect on the angle. The tilting must therefore be caused by molecular packing (i.e. crystal) forces. This twisting also occurs in the crystal structure of potassium hydrogen di-p-chlorobenzoate (Mills and Speakman, 1963) where the angle of twist is 9° . As can be seen from Fig. 2 and Fig. 3 the benzene rings themselves are very tightly packed, and this close packing of the benzene rings is in conflict with the demands of the packing of the

oxygen atoms about the potassium ion.

The plane through C(5), O(3) and C(8) is:

$$0.1899X + 0.8231Y - 0.5352Z = 1.0544, \quad (3)$$

and this makes an angle of 4° with the plane of the benzene ring, and an angle of 11.2° with the plane of the carboxyl group.

An unusual, but not unique feature of this structure is that the "hydrogen" of the hydrogen bond is not located at a centre of inversion as is usual in "Type A" acid salts, but is situated on a diad axis. Another example of a symmetrical hydrogen bond in this situation is given by Mills and Speakman (1961) in the structure of sodium hydrogen diacetate. The hydrogen bond itself is of the short symmetrical type and has a length (OO) of $2.476 \pm 7 \text{ \AA}$. The low value of the estimated standard deviation is obtained though the hydrogen bond lies across a diad axis. The estimated standard deviation of the x-coordinate of the O(1) atom is very low, and the standard deviation of the y-coordinate has no effect.

The packing of the oxygen atoms round the potassium ion is best described as a distorted octahedron. The potassium ion is situated on a diad axis, surrounded by six oxygen atoms, all at distances in the range 2.74 to 2.97 \AA , and the closest contacts are between the "double-bonded" oxygen atoms and the potassium ion.

TABLE 4.1

FRACTIONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
K(1)	0.0000 ± 0	-0.2059 ± 5	0.2500 ± 0
O(1)	0.0309 ± 1	0.3523 ± 9	0.3407 ± 9
O(2)	0.0450 ± 1	0.1226 ± 9	0.0973 ± 9
O(3)	0.2049 ± 1	0.2412 ± 10	0.5271 ± 9
C(1)	0.0543 ± 2	0.2355 ± 14	0.2429 ± 13
C(2)	0.0936 ± 2	0.2428 ± 10	0.3246 ± 13
C(3)	0.1030 ± 2	0.3332 ± 12	0.5139 ± 12
C(4)	0.1397 ± 2	0.3341 ± 11	0.5935 ± 13
C(5)	0.1680 ± 2	0.2490 ± 12	0.4648 ± 14
C(6)	0.1588 ± 2	0.1619 ± 12	0.2738 ± 13
C(7)	0.1227 ± 2	0.1553 ± 13	0.1984 ± 14
C(8)	0.2159 ± 20	0.3381 ± 19	0.7202 ± 17

TABLE 4.2

ATOMIC COORDINATES AND E.S.D.
IN ANGSTROMS

ATOM	X	Y	Z
K(1)	0.000 ± 0	-1.452 ± 3	1.539 ± 0
O(1)	1.105 ± 4	2.483 ± 6	2.098 ± 6
O(2)	1.609 ± 5	0.864 ± 7	0.599 ± 6
O(3)	7.330 ± 4	1.700 ± 7	3.246 ± 6
C(1)	1.943 ± 7	1.660 ± 10	1.496 ± 8
C(2)	3.350 ± 6	1.711 ± 7	1.999 ± 8
C(3)	3.684 ± 6	2.349 ± 9	3.165 ± 8
C(4)	4.996 ± 6	2.355 ± 8	3.655 ± 8
C(5)	6.009 ± 5	1.755 ± 9	2.862 ± 8
C(6)	5.681 ± 6	1.141 ± 9	1.686 ± 8
C(7)	4.388 ± 6	1.095 ± 9	1.222 ± 8
C(8)	7.723 ± 72	2.383 ± 14	4.435 ± 11

TABLE 4.3

FRACTIONAL COORDINATES
AND TEMPERATURE FACTORS
FOR HYDROGEN ATOMS

ATOM	X	Y	Z	U
H(3)	0.081	0.407	0.608	0.0658
H(4)	0.148	0.403	0.743	0.0454
H(6)	0.179	0.081	0.178	0.0714
H(7)	0.116	0.076	0.033	0.0446

TABLE 4.4

ANISOTROPIC TEMPERATURE FACTORS
AND THEIR E.S.D.

ATOM	U11	U22	U33	2U23	2U31	2U12
K(1)	0.0447 11	0.0570 17	0.0404 17	0.0000 0	-0.0116 18	0.0000 0
O(1)	0.0262 19	0.0636 37	0.0564 34	-0.0193 63	-0.0116 39	-0.0001 45
O(2)	0.0415 22	0.0661 41	0.0532 36	-0.0143 68	-0.0104 43	-0.0063 54
O(3)	0.0262 19	0.0674 41	0.0642 37	-0.0073 67	0.0010 40	0.0091 47
C(1)	0.0387 33	0.0568 55	0.0415 52	-0.0033 87	0.0041 56	-0.0041 74
C(2)	0.0283 27	0.0310 38	0.0396 44	0.0173 71	0.0117 53	-0.0005 53
C(3)	0.0238 25	0.0523 50	0.0431 48	-0.0150 79	0.0025 51	0.0095 60
C(4)	0.0304 27	0.0436 45	0.0398 46	-0.0001 75	0.0088 53	0.0061 61
C(5)	0.0178 24	0.0494 48	0.0591 54	0.0130 82	-0.0004 52	-0.0022 58
C(6)	0.0299 27	0.0422 47	0.0457 50	0.0016 76	0.0148 54	-0.0090 63
C(7)	0.0321 29	0.0494 50	0.0443 48	-0.0044 86	0.0132 58	-0.0052 66
C(8)	0.0278 29	0.1053 90	0.0822 73	-0.0410 134	-0.0436 70	0.0068 91

TABLE 4.5

PRINCIPAL VALUES AND DIRECTIONS OF
VIBRATION TENSORS

ATOM	^{o2} U A	D1	D2	D3
K(1)	0.0364	0.5711	0.0000	0.8208
	0.0570	1.0000	0.0000	0.0000
	0.0487	-0.8208	0.0000	0.5711
O(1)	0.0250	0.9792	0.0505	0.1966
	0.0705	-0.0753	-0.8091	0.5828
	0.0506	0.1885	-0.5855	-0.7885
O(2)	0.0693	0.0295	-0.9171	0.3977
	0.0384	0.8680	0.2067	0.4108
	0.0531	-0.4589	0.3410	0.8204
O(3)	0.0257	-0.9936	0.1104	0.0234
	0.0701	-0.0822	-0.8497	0.5208
	0.0620	0.0774	0.5155	0.8534
C(1)	0.0573	0.1220	-0.9854	0.1191
	0.0376	-0.8954	-0.0575	0.4415
	0.0422	-0.4282	-0.1605	-0.8893
C(2)	0.0463	0.2667	0.4710	0.8408
	0.0230	-0.6129	-0.5904	0.5251
	0.0296	0.7438	-0.6554	0.1312
C(3)	0.0569	-0.1083	-0.8773	0.4676
	0.0227	-0.9734	0.1890	0.1292
	0.0396	-0.2017	-0.4411	-0.8745
C(4)	0.0281	-0.9197	0.1825	0.3476
	0.0445	0.2787	0.9270	0.2509
	0.0412	-0.2764	0.3277	-0.9034
C(5)	0.0178	0.9994	-0.0007	0.0349
	0.0624	-0.0151	0.4490	0.8934
	0.0462	0.0315	-0.8928	0.4493
C(6)	0.0258	-0.9010	-0.2633	0.3448
	0.0488	0.3930	-0.1589	0.9057
	0.0433	-0.1837	0.9515	0.2467

c(7)	0.0291	-0.9186	-0.0756	0.3879
	0.0516	0.2779	-0.8215	0.4979
	0.0451	0.2810	0.5652	0.7756

c(8)	0.0200	0.9381	0.0452	0.3435
	0.1195	-0.1600	-0.8229	0.5452
	0.0758	0.3073	-0.5664	-0.7647

TABLE 4.6

Bond-lengths and E.S.D.

O(1) - C(1)	1.320 \pm 10	
O(2) - C(1)	1.244 \pm 11	
C(1) - C(2)	1.495 \pm 9	
C(2) - C(3)	1.370 \pm 11	} Benzene Ring
C(3) - C(4)	1.401 \pm 9	
C(4) - C(5)	1.411 \pm 10	
C(5) - C(6)	1.367 \pm 12	
C(6) - C(7)	1.374 \pm 9	
C(7) - C(2)	1.436 \pm 10	
C(5) - O(3)	1.376 \pm 7	
C(3) - C(8)	1.427 \pm 23	

TABLE 4.7

Final observed and calculated structure factors.

H	K	L	F	DBS	F	CA	H	K	L	F	DBS	F	CA	H	K	L	F	DBS	F	CA	H	K	L	F	DBS	F	CA	H	K	L	F	DBS	F	CA
0	0	4	56.9	51.1	3	6	2	10.7	9.2	7	1	1	26.8	27.9	10	1	2	28.9	28.0	13	1	3	8.8	20.5	18	2	3	10.1	20.9					
0	0	6	14.1	14.7	3	6	3	5.4	6.6	7	1	1	35.2	37.5	10	1	2	21.1	24.0	13	1	4	14.0	15.3	18	2	4	14.0	16.0	18.9	18.9			
0	2	1	72.4	72.4	2	7	0	14.7	13.5	7	2	47.2	39.8	10	1	3	6.8	7.3	13	1	4	10.3	19.3	18	3	1	8.3	6.3	17.4	17.4				
0	2	2	33.7	33.7	3	2	2	3.5	4.2	7	1	2	31.3	29.0	10	1	3	6.5	7.3	13	1	5	11.1	2.6	18	3	2	19.7	19.7	17.4	17.4			
0	2	3	25.6	25.6	4	0	2	4.7	5.4	7	1	3	17.1	21.3	10	1	4	12.8	11.5	13	2	1	10.5	8.6	18	3	1	8.1	8.1	8.9	8.9			
0	4	1	8.8	8.8	4	0	4	14.8	16.5	7	2	1	28.4	22.7	10	1	5	7.6	6.9	13	2	2	43.3	41.7	18	4	3	14.3	14.3	14.3	14.3			
0	4	2	4.1	4.2	4	1	1	87.6	76.7	7	1	4	22.8	22.7	10	1	7	4.1	3.9	13	2	4	2.7	1.1	18	4	1	18.6	19.7	17.4	17.4			
0	4	3	9.5	8.8	4	1	2	5.5	45.9	7	1	5	5.5	9.1	10	2	0	86.8	63.5	13	3	0	22.2	25.6	18	4	1	23.3	18.9	18.9	18.9			
0	4	4	33.2	27.0	4	1	2	49.7	45.9	7	2	7	18.8	18.0	10	2	1	18.0	11.8	13	3	2	43.3	41.7	18	4	2	20.8	20.5	20.5	20.5			
0	6	2	13.8	9.6	4	3	3	46.7	38.2	7	2	1	11.3	11.3	10	2	3	5.4	4.0	13	3	3	37.0	32.6	18	5	2	9.6	6.3	6.3	6.3			
0	6	3	14.7	4.8	4	4	1	4	5.7	3.0	7	2	28.8	29.8	10	2	4	25.0	21.0	13	4	4	21.3	22.0	18	5	4	6.5	8.8	8.8	8.8			
0	6	4	1.7	4.0	4	5	5	5.5	3.0	7	2	4	9.2	7.5	10	3	0	33.4	22.9	13	5	2	27.4	23.6	18	6	3	2.5	3.0	3.0	3.0			
0	8	2	7.2	2.2	4	1	6	4.4	8.4	7	3	1	44.2	45.0	10	3	3	2.4	4.7	13	6	4	15.2	13.0	18	6	2	3.0	2.0	2.0	2.0			
1	0	2	24.4	47.8	4	1	7	4.4	8.4	7	3	2	43.2	39.2	10	3	4	42.8	16.3	13	6	4	4.5	4.2	18	6	3	10.7	10.0	10.0	10.0			
1	0	4	14.4	17.2	4	2	0	14.5	8.5	7	3	3	93.5	42.3	10	4	0	53.1	55.5	13	7	0	27.1	28.9	18	7	2	2.5	3.1	3.1	3.1			
1	0	8	2.7	2.7	4	2	2	36.7	34.5	7	3	4	13.2	11.1	10	4	2	2.7	2.8	13	8	3	14.3	14.9	18	7	1	9.9	7.9	7.9	7.9			
1	1	0	24.7	23.6	4	3	3	22.2	20.2	7	4	1	14.2	12.1	10	4	2	13.5	19.7	13	9	4	7.8	9.4	18	8	0	4.7	7.1	7.1	7.1			
1	1	1	8.2	7.4	4	4	2	28.3	24.1	7	4	2	14.2	12.1	10	4	4	3.8	7.5	13	10	4	2	3.1	6.7	18	9	0	14.7	13.0	13.0	13.0		
1	1	2	17.3	21.8	4	3	2	43.0	40.5	7	4	3	8.5	7.1	10	5	2	3.0	2.5	13	11	3	3.2	3.1	18	10	4	63.8	59.8	59.8	59.8			
1	1	3	13.4	14.7	4	3	3	18.9	20.2	7	4	4	27.7	26.2	10	5	3	2.7	1.4	13	12	0	4.8	3.4	18	11	1	24.6	21.7	21.7	21.7			
1	1	3	13.4	16.1	4	3	4	9.9	7.9	7	5	2	14.4	14.5	10	5	4	3.2	3.4	13	13	2	2.7	2.7	18	11	1	21.3	20.7	20.7	20.7			
1	1	4	2.4	2.4	4	4	0	10.4	13.4	7	5	2	24.2	22.5	10	6	0	9.7	10.8	14	14	0	14.1	11.2	18	12	1	22.7	22.9	22.9	22.9			
1	1	5	24.3	25.6	4	4	2	9.4	8.1	7	5	3	30.0	21.8	10	6	1	4.0	3.1	14	15	2	2.2	2.2	18	13	2	22.7	22.9	22.9	22.9			
1	1	6	13.5	12.4	4	4	3	16.2	13.4	7	5	4	17.1	13.4	10	6	2	3.1	2.2	14	16	0	51.9	52.1	18	13	3	31.2	31.1	31.1	31.1			
1	1	7	15.9	17.9	4	4	4	7.1	7.0	7	5	3	2.9	1.9	10	7	2	2.8	4.4	14	17	0	51.1	52.1	18	13	4	28.4	31.1	31.1	31.1			
1	1	8	12.8	9.4	4	5	3	16.2	16.4	7	7	0	4.8	2.9	10	8	0	2.7	2.9	14	18	0	4.4	3.7	18	14	5	13.6	11.4	11.4	11.4			
1	2	2	1.4	1.0	4	5	3	10.2	10.4	7	7	0	2.9	0.6	11	0	0	13.2	9.4	14	19	0	6.6	3.0	18	15	2	27.1	27.6	27.6	27.6			
1	2	3	13.0	13.1	4	5	4	14.4	13.5	7	7	0	13.2	9.4	11	0	4	7.7	12.6	14	20	1	11.1	11.1	18	16	2	20.2	19.2	19.2	19.2			
1	2	4	2.4	2.4	4	6	5	5.4	6.2	7	8	0	3.8	2.2	11	0	6	27.8	22.7	14	1	1	36.3	35.4	18	17	4	3.0	1.8	1.8	1.8			
1	3	1	21.4	17.2	4	7	2	2.9	8.2	7	8	0	13.2	14.3	11	1	8	13.2	10.2	14	2	1	25.4	25.3	18	18	2	4.2	3.1	3.1	3.1			
1	3	2	24.9	49.7	4	7	2	2.9	8.2	7	8	0	13.2	14.3	11	1	1	58.7	37.1	14	3	3	34.2	41.9	18	19	3	12.1	13.3	13.3	13.3			
1	3	3	5.6	4.6	4	8	0	2.0	3.5	7	8	0	6	11.2	6.7	11	1	1	91.9	59.1	14	4	4	41.1	41.9	18	20	4	33.1	33.1	33.1	33.1		
1	3	4	23.7	23.7	4	8	2	14.9	8.9	7	8	1	1	10.4	10.4	11	1	2	4.1	4.1	14	5	4	2.8	13.7	18	21	3	18.4	18.7	18.7	18.7		
1	4	2	23.9	17.1	5	0	4	22.5	22.7	7	9	1	1	29.1	24.3	11	2	4	4.1	21.2	14	6	1	19.3	16.8	18	22	4	3.9	3.4	3.4	3.4		
1	4	3	2.4	1.0	5	0	6	1.4	1.4	7	9	1	2	28.7	30.3	11	3	3	39.5	28.0	14	7	3	35.8	35.9	18	23	4	1.9	3.4	3.4	3.4		
1	4	4	1.4	1.4	5	1	1	10.4	10.4	7	9	1	3	26.3	21.9	11	4	4	8.4	6.0	14	8	2	13.4	14.1	18	24	4	13.8	13.9	13.9	13.9		
1	4	5	24.9	17.8	5	1	2	30.3	24.9	7	9	1	3	26.3	21.9	11	4	5	8.4	6.0	14	9	2	25.2	24.3	18	25	4	10.7	10.7	10.7	10.7		
1	5	0	36.0	36.0	5	1	2	30.3	24.9	7	9	1	4	8.5	4.8	11	5	4	8.4	6.0	14	10	2	2.1	1.5	18	26	4	7.9	7.9	7.9	7.9		
1	5	3	2.7	1.4	5	1	2	33.0	24.5	7	9	1	4	8.5	4.8	11	7	4	12.7	14.0	14	11	4	23.4	21.7	18	27	5	17.1	14.7	14.7	14.7		
1	5	4	23.2	20.5	5	1	4	9.2	10.5	7	9	1	5	5.5	5.6	11	2	1	33.5	24.6	14	12	1	10.3	12.9	18	28	2	19.9	15.5	15.5	15.5		
1	6	2	1.4	1.4	5	2	1	14.4	14.4	7	9	2	1	14.4	14.4	11	2	2	12.2	12.5	14	13	2	15.5	12.9	18	29	4	3.0	3.0	3.0	3.0		
1	6	3	2.7	2.6	5	2	1	6.8	10.0	7	9	2	1	4.6	3.5	11	2	3	12.2	12.5	14	14	3	10.7	8.2	18	30	4	9.7	8.4	8.4	8.4		
1	6	4	13.6	16.7	5	2	5	23.2	20.4	7	9	2	2	4.0	0.1	11	3	4	14.1	12.0	14	15	4	4.3	6.9	18	31	2	3.0	1.4	1.4	1.4		
1	7	2	2.9	2.9	5	3	6	5.5	1.1	7	9	2	3	4.2	3.5	11	3	5	4.2	3.5	14	16	4	37.8	32.4	18	32	3	3.5	1.9	1.9	1.9		
1	7	3	13.2	14.4	5	3	4	25.1	24.4	7	9	2	4	1.4	1.4	11	4	6	2.8	2.8	14	17	5	18.2	19.5	18	33	4	4.4	2.1	2.1	2.1		
1	7	4	14.1	14.8	5	3	2	12.1	9.0	7	9	2	5	13.8	11.1	11	4	7	1.1	1.1	14	18	6	2.8	0.1	18	34	2	7.2	2.5	2.5	2.5		
1	7	5	14.1	14.6	5	3	3	29.3	25.4	7	9	2	6	3.2	1.1	11	5	8	3.2	3.2	14	19	7	17.1	15.7	18	35	0	52.7	43.0	43.0	43.0		
1	8	1	12.7	12.7	5	4	6	1.4	1.4	7	9	3	1	12.7	12.7	11	6	9	4.1	3.4	14	20	8	4.4	4.8	18	36	0	69.7	43.0	43.0	43.0		
1	8	2	15.1	15.1	5	4	7	1.4	1.4	7	9	3	2	12.7	12.7	11	7	10	10.2	10.2	14	21	9	2.1	2.1	18	37	4	41.1	41.1	41.1	41.1		
1	8	3	12.7	12.7	5	4	8	1.4	1.4	7	9	3	3	12.7	12.7	11	8	11	12.7	12.7	14	22	10	2.1	2.1	18	38	4	36.7	41.6	41.6	41.6		
1	8	4	15.1	15.1	5	4	9	1.4	1.4	7	9	3	4	12.7	12.7	11	9	12	12.7	12.7	14	23												

Tables of Bond-Lengths, Bond-Angles and Interatomic

Contacts

In the following tables, coordinates of atoms (x, y, z) denoted only by a bracketed number are those given in TABLE 4.1. The others are related to this set as follows:

A	x, -y,	$\frac{1}{2} + z$
B	-x, -y,	-z
C	x, -y,	$-\frac{1}{2} + z$
D	-x, -y,	1 - z
E	-x, +y,	$\frac{1}{2} - z$
F	-x, -1+y,	$\frac{1}{2} - z$
G	x, -1+y,	z

TABLE 4.8

Potassium-Oxygen Contacts

K(1)	O(2A)	2.740
K(1)	O(2B)	2.740
K(1)	O(1C)	2.939
K(1)	O(1D)	2.939
K(1)	O(2)	2.973
K(1)	O(2E)	2.973
K(1)	O(1F)	3.351
K(1)	O(1G)	3.351

TABLE 4.9

Oxygen-Oxygen Contact

O(1) O(1E) 2.476 \pm 15

TABLE 4.10

Bond-Angles and E.S.D.

O(1) - C(1) - O(2)	123.8 \pm 7
O(1) - C(1) - C(2)	115.0 \pm 7
O(2) - C(1) - C(2)	121.1 \pm 7
C(1) - C(2) - C(3)	122.1 \pm 6
C(1) - C(2) - C(7)	118.9 \pm 7
C(2) - C(3) - C(4)	121.9 \pm 7
C(3) - C(4) - C(5)	118.1 \pm 7
C(4) - C(5) - C(6)	119.9 \pm 6
C(5) - C(6) - C(7)	122.1 \pm 7
C(6) - C(7) - C(2)	118.9 \pm 7
C(4) - C(5) - O(3)	123.1 \pm 7
C(6) - C(5) - O(3)	116.9 \pm 6
C(5) - O(3) - C(8)	118.5 \pm 2.8

TABLE 4.11
MEAN PLANES

Plane 1

$$O(1) - O(2) - C(1) - C(2)$$

$$0.2102 X + 0.6987 Y - 0.6838 Z = 0.5361 (\text{\AA})$$

$$\text{Root Mean Square Deviation from Plane} = 0.0054 (\text{\AA})$$

$$\text{Maximum Deviation (C(1))} = 0.009 (\text{\AA})$$

Plane 2

$$C(1) - C(2) - C(3) - C(4) - C(5) - C(6) - C(7) - O(3)$$

$$0.1516 X + 0.8592 Y - 0.4886 Z = 1.0063 (\text{\AA})$$

$$\text{Root Mean Square Deviation from Plane} = 0.0147 (\text{\AA})$$

$$\text{Maximum Deviation (C(3))} = 0.024 (\text{\AA})$$

Plane 3

$$O(3) - C(5) - C(8)$$

$$0.1899 X + 0.8231 Y - 0.5352 Z = 1.0544 (\text{\AA})$$

Plane 4

$$C(2) - C(3) - C(4) - C(5) - C(6) - C(7)$$

$$0.1501 X + 0.8594 Y - 0.4887 Z = 1.0054 (\text{\AA})$$

$$\text{Root Mean Square Deviation from Plane} = 0.0117 (\text{\AA})$$

$$\text{Maximum Deviation (C(3))} = 0.019 (\text{\AA})$$

Deviations of atoms not in the Plane:

$$O(3) = -0.031 (\text{\AA})$$

$$C(1) = -0.018 (\text{\AA})$$

$$C(8) = 0.034 (\text{\AA})$$

ANGLES

Plane 1 - Plane 2	=	14.90°
Plane 1 - Plane 3	=	11.18°
Plane 1 - Plane 4	=	14.92°
Plane 2 - Plane 3	=	4.03°
Plane 2 - Plane 4	=	0.09°
Plane 3 - Plane 4	=	4.08°

TABLE 4.12

BOND-LENGTHS OF THE $\begin{array}{c} \text{O} \\ \parallel \\ \text{C}(1) - \text{C}(2) \\ \diagup \\ \text{O} \end{array}$ BOND IN VARIOUS ACID SALTS

C(1) - C(2) (Å)	COMPOUND
1.492	potassium hydrogen diacetate
1.486	potassium hydrogen di(p-chloro-benzoate)
1.515	potassium hydrogen di(tri-fluoro-acetate)
1.52	caesium hydrogen di(tri-fluoro-acetate)
1.465	potassium hydrogen di crotonate

Average = 1.495

References

Mills and Speakman (1961)

Mills and Speakman (1963)

Golič and Speakman (1965B)

(The structure of potassium hydrogen di crotonate is described in Chapter 5)

CHAPTER 5

THE STRUCTURE DETERMINATION OF POTASSIUM HYDROGEN DICROTONATE

Contents

- 5.1. Introduction
- 5.2. Experimental
- 5.3. Choice of Space Groups
- 5.4. Structure Solution
- 5.5. Least-Squares Refinement
- 5.6. Discussion
- 5.7. Appendix: Cell Dimensions and Reduced Cell.

Chapter 5

THE STRUCTURE DETERMINATION

OF

POTASSIUM HYDROGEN DICROTONATE

5.1. Introduction

The crystal structures of the acid salts of monocarboxylic acids have been intensively studied over the past 20 years by Speakman and his co-workers. Their interest has centred on the very short hydrogen bonds which occur in some of these compounds. The compounds have been assigned to two classes (See Chapter 4) on the basis of their crystal structure as revealed by infra-red spectroscopy, X-ray diffraction and neutron diffraction.

Prior to the present work, the crystal structure of crotonic acid itself $\text{CH}_3 - \text{CH} = \text{CH} - \text{CO}_2\text{H}$ was determined by S. A. Sutherland in 1961. Later, crystals of potassium hydrogen dicrotonate, rubidium hydrogen dicrotonate and ammonium hydrogen dicrotonate were prepared by I. McCrorie (B.Sc. Thesis, Glasgow, 1963) who measured the cell dimensions of the potassium compound, and derived the "reduced" cell from the morphological unit cell by application of the Delaunay reduction.

This chapter describes the structure analysis and the crystal structure of potassium hydrogen dicrotonate from three-dimensional X-ray diffraction data.

5.2. Experimental

Preparation of crystals: potassium hydrogen dicrotonate is readily prepared by dissolving crotonic acid with potassium hydroxide (in the ratio of two equivalents to one equivalent) in a little 50% aqueous ethanol, and allowing the solution to cool.

Crystal Data

Potassium hydrogen dicrotonate KH $(C_4H_5O_2)_2$;

M = 210.3; triclinic; a = 12.46;

b = 6.02; c = 7.45 Å; $\alpha = 66^\circ 31'$;

$\beta = 103^\circ 39'$; $\gamma = 95^\circ 12'$.

The reduced cell is given in the Appendix to this chapter, along with the matrix for transforming the above cell to the reduced cell.

The unit cell given above was used throughout these investigations. All atomic fractional co-ordinates and indexing of reflexions refer to it and its corresponding reciprocal cell, not to the reduced cell.

V = 496.72 Å³; Z = 2; D_m = 1.37 g/cc;

D_c = 1.40 g/cc; space group P $\bar{1}$ (See below);

F₀₀₀ = 220;

Absorption coefficient (CuK α radiation) $\mu = 46 \text{ cm}^{-1}$.

5.3. Choice of Space Groups

The space groups $P1$ and $P\bar{1}$ cannot be distinguished by systematic absences of classes of X-ray reflexions. Initially an attempt was made to distinguish between the two by applying the "N(Z)" and "variance" statistical tests to the distribution of intensities. Both these simple tests indicated that the space group was $P1$. It has been shown by Sim (1959) and Foster and Hargreaves (1963) that the presence of heavy atoms can distort the intensity distribution. Indeed it was found that the N(Z) distribution was almost identical to one given by Cochran (1963) for a heavy-atom compound in the space group $P\bar{1}$.

The Patterson map was interpreted in terms of a centrosymmetric structure and this has been confirmed by the successful refinement of the structure in $P\bar{1}$, and by comparing the results of the refinement with those of a parallel refinement in the space group $P1$.

5.4. Structure Solution

The intensities of 1225 independent X-ray reflexions were estimated visually from multiple-film Weissenberg photographs of the hko , okl , hol , hll - and $h2l$ reciprocal lattice nets. No absorption corrections were applied. The position of the potassium-potassium vector peak could not be located unambiguously from the two (uvo and vow) short-axial projections of the Patterson function. Two three-dimensional Patterson maps were therefore calculated, one "sharpened", the other unsharpened. An empirical sharpening function, which increased the magnitudes of the F^2 coefficients of reflexions as $\sin \theta$ increased, was used.

It was then realised that the shortest distance between potassium ions in the crystal could not be less than $\sim 3\text{\AA}$, and was probably in the range $3.1\text{\AA} - 3.4\text{\AA}$. When looking for the potassium-potassium vector peak, therefore, only peaks at least 3\AA from the origin were considered. This left only one major peak, from which the co-ordinates of the potassium atom were derived.

Structure factors for all the reflexions calculated for this one potassium atom alone gave an R-factor of 58%. An electron-density map was then calculated using the observed structure amplitudes with the phases of the

structure-factors calculated for the potassium atom alone. This revealed an imperfect image of the structure from which it was possible to pick out the atomic positions of the four oxygen atoms.

After four cycles of structure-factors and electron-density maps, the positions of all the atoms, except hydrogen, were known and the R-factor was now 31%.

5.5. Least-Squares Refinement

The structure was now refined by the Least-Squares method. Only co-ordinate parameters and a single isotropic temperature-parameter were refined for each atom. A single over-all scale-parameter was used, and the observations were given equal weights, but no "unobserved" reflexions were included in the refinement at this stage. Four cycles of refinement brought the R-factor to 18%, with considerable shifts in atomic co-ordinates and temperature parameters.

Unobserved reflexions were included at $\frac{1}{2}$ I min (i.e. $\frac{1}{4}|F_o|/\text{local average}$), and a weighting scheme was fitted by means of the program described in Chapter 2. The weighting scheme was:

$$W = K \exp - (A + B F_o + C F_o^2 + D \lambda^2 / \sin^2 \theta + E \lambda^4 / \sin^4 \theta + F \sin^2 \theta / \lambda^2)$$

where K is an arbitrary scale factor and A, B, C, D, E and F are fitted by the Least-Squares method, so that

$$W = 1 / \langle \Delta^2 \rangle \text{ where } \langle \Delta^2 \rangle \text{ is the local average of } (F_o - F_c)^2.$$

Each atom's isotropic temperature parameter was replaced by 6 anisotropic temperature parameters, and each reciprocal net was given its own scale factor.

Three cycles of Least-Squares reduced the R-factor

to 15.5%. It was then apparent that the weighting scheme would have to be re-adjusted to maintain a "flat" distribution of $\langle W\Delta^2 \rangle$. Anisotropic structure factors were calculated, using R. Pollard's (1966) program. These were now used for the calculation of an $(F_o - F_c)$ map and to obtain re-adjusted weighting parameters.

The difference map revealed a complex depressed region surrounding the heavy-atom site, which was thought to have been caused by uncorrected absorption and extinction errors. Four peaks were located near the positions where it was expected the hydrogen atoms of the double-bonded carbon atoms would lie, but no definite peaks were observed which could be assigned to the hydrogen atoms of the methyl groups. The acidic hydrogen atom could not be located.

Least-Squares Refinement was then continued for another three cycles. Although the final R-factor over all the reflexions is 15.2%, it can be seen from the tables of the structure factors that the agreement between the observed and calculated values of individual structure factors is satisfactory. Approximately 30% of the reflexions were "unobserved" and these contribute significantly to the R-factor.

To check that the space-group was really $P\bar{1}$ and was

FIG. 1

Bond-lengths

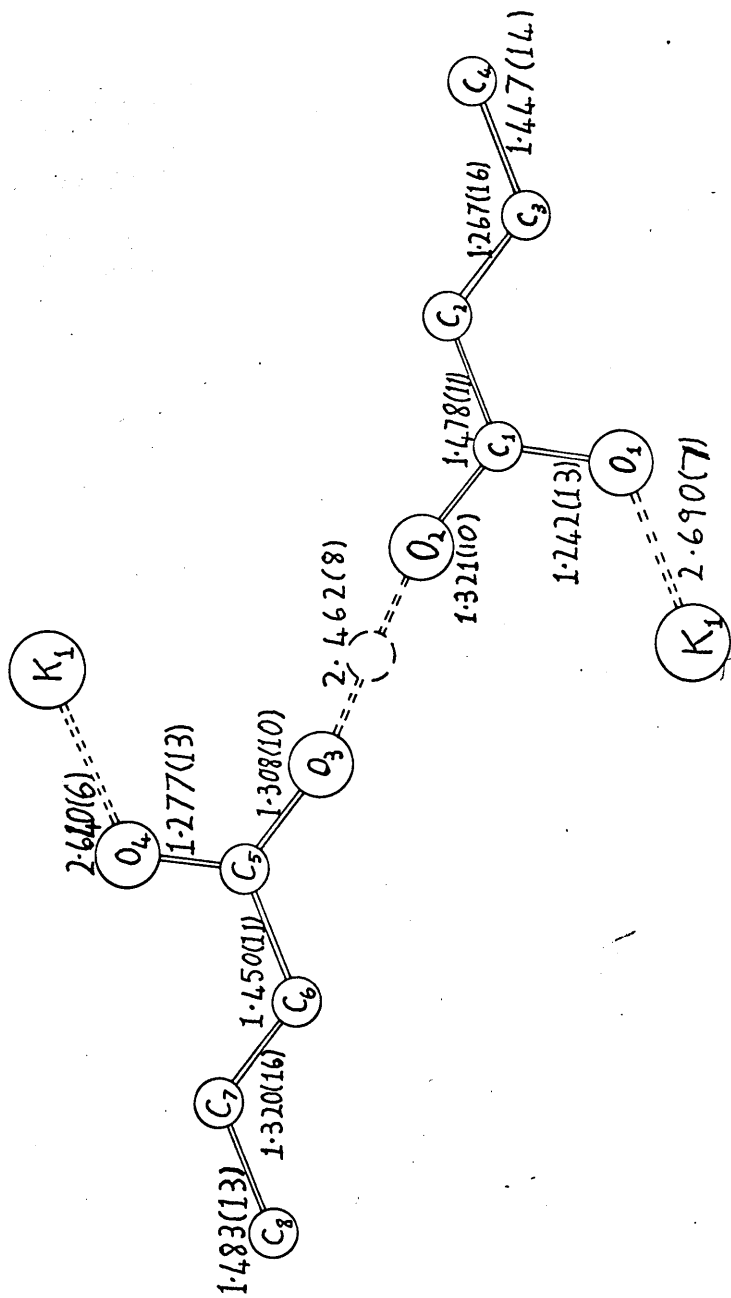


FIG. 2
Bond-angles

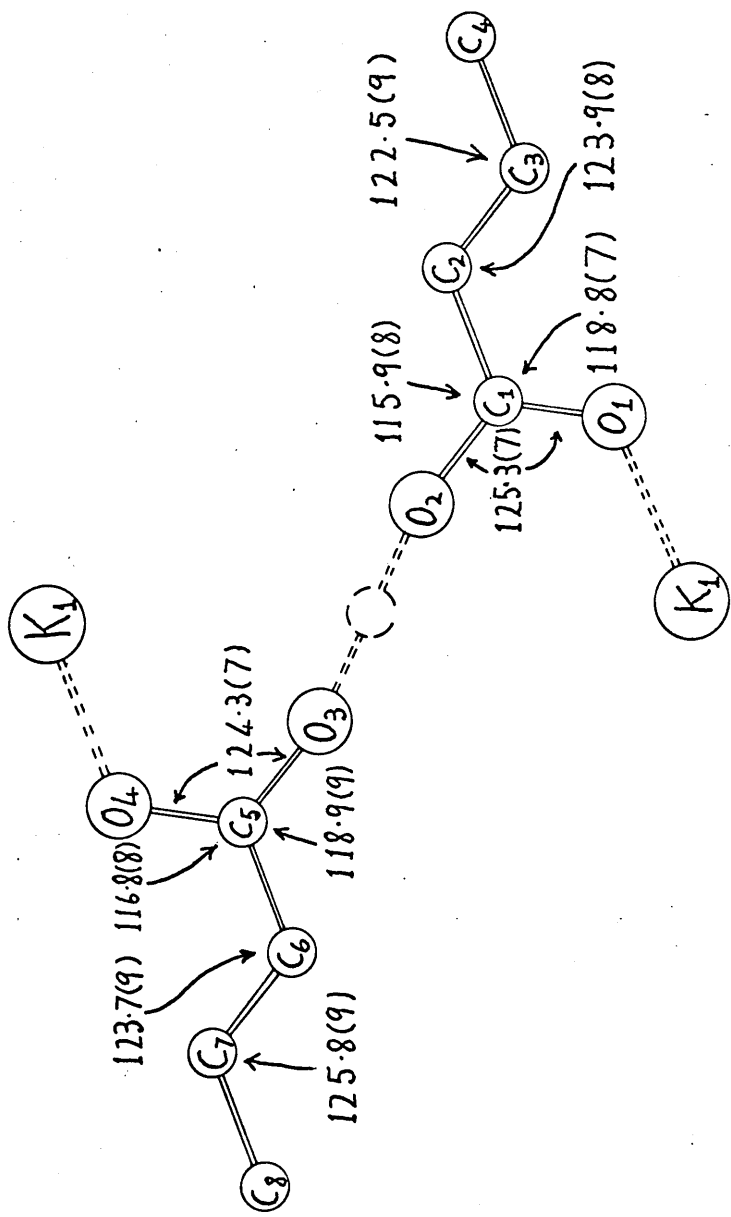


TABLE 5.1

FRACTIONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
K(1)	-0.0900 ± 1	-0.0390 ± 4	0.2143 ± 3
O(1)	0.7356 ± 5	0.2002 ± 15	0.2054 ± 11
O(2)	0.8198 ± 4	0.4924 ± 13	0.3107 ± 10
O(3)	0.0001 ± 4	0.3243 ± 12	0.3464 ± 8
O(4)	0.0476 ± 4	0.7052 ± 13	0.1510 ± 8
C(1)	0.7326 ± 5	0.3775 ± 19	0.2498 ± 11
C(2)	0.6240 ± 6	0.4774 ± 23	0.2276 ± 13
C(3)	0.6115 ± 6	0.6527 ± 23	0.2741 ± 14
C(4)	0.5038 ± 8	0.7462 ± 27	0.2503 ± 19
C(5)	0.0662 ± 5	0.4780 ± 20	0.2439 ± 11
C(6)	0.1665 ± 6	0.3879 ± 21	0.2290 ± 13
C(7)	0.1929 ± 5	0.1561 ± 20	0.3044 ± 12
C(8)	0.2912 ± 8	0.0540 ± 27	0.2842 ± 17

TABLE 5.2

COORDINATES AND E.S.D. IN ANGSTROMS
WITH RESPECT TO THE CRYSTAL AXES

ATOM	X	Y	Z
K(1)	-1.121 \pm 1	-0.235 \pm 2	1.597 \pm 2
O(1)	9.165 \pm 6	1.205 \pm 9	1.531 \pm 8
O(2)	10.215 \pm 5	2.964 \pm 8	2.315 \pm 8
O(3)	0.002 \pm 5	1.953 \pm 7	2.581 \pm 6
O(4)	0.593 \pm 5	4.246 \pm 8	1.125 \pm 6
C(1)	9.129 \pm 6	2.272 \pm 12	1.861 \pm 8
C(2)	7.775 \pm 8	2.874 \pm 14	1.696 \pm 9
C(3)	7.620 \pm 8	3.929 \pm 14	2.042 \pm 11
C(4)	6.277 \pm 10	4.492 \pm 16	1.865 \pm 14
C(5)	0.825 \pm 7	2.877 \pm 12	1.817 \pm 8
C(6)	2.074 \pm 7	2.335 \pm 13	1.706 \pm 10
C(7)	2.403 \pm 6	0.940 \pm 12	2.268 \pm 9
C(8)	3.629 \pm 9	0.325 \pm 16	2.118 \pm 13

ORTHOGONAL COORDINATES

The orthogonal axes (X, Y, Z) are defined as follows:

X is parallel to a^*

Z is parallel to c

Y is normal to X and Z, so as to
complete a right-handed set.

TABLE 5.3

ORTHOGONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
K(1)	-1.089 ± 1	-0.220 ± 2	1.768 ± 2
O(1)	8.906 ± 6	1.140 ± 8	-0.152 ± 7
O(2)	9.926 ± 5	2.757 ± 7	1.085 ± 7
O(3)	0.002 ± 5	1.791 ± 7	3.358 ± 5
O(4)	0.576 ± 5	3.896 ± 7	2.677 ± 6
C(1)	8.871 ± 6	2.118 ± 11	0.612 ± 7
C(2)	7.556 ± 8	2.665 ± 13	1.006 ± 8
C(3)	7.404 ± 8	3.632 ± 13	1.810 ± 10
C(4)	6.100 ± 10	4.143 ± 15	2.173 ± 13
C(5)	0.802 ± 7	2.642 ± 11	2.769 ± 7
C(6)	2.016 ± 7	2.149 ± 12	2.147 ± 9
C(7)	2.335 ± 6	0.871 ± 11	2.076 ± 8
C(8)	3.526 ± 9	0.312 ± 15	1.391 ± 11

TABLE 5.4

FRACTIONAL COORDINATES AND TEMPERATURE FACTORS
FOR HYDROGEN ATOMS

ATOM	X	Y	Z	U
H(1)	0.551	0.384	0.168	0.0500
H(2)	0.684	0.755	0.323	0.0500
H(3)	0.225	0.482	0.178	0.0500
H(4)	0.134	0.071	0.363	0.0500

TABLE 5.5

ORTHOGONAL COORDINATES FOR HYDROGEN ATOMS

ATOM	X	Y	Z
H(1)	6.67	2.14	0.55
H(2)	8.29	4.20	2.20
H(3)	2.73	2.67	1.82
H(4)	1.63	0.40	2.48

TABLE 5.6

ANISOTROPIC TEMPERATURE FACTORS AND E.S.D.

ATOM	U11	U22	U33	2U23	2U31	2U12
K(1)	0.0295 7	0.0212 12	0.0369 9	-0.0251 20	0.0305 14	0.0063 19
O(1)	0.0356 27	0.0505 54	0.0587 38	-0.0522 87	0.0081 54	-0.0020 71
O(2)	0.0263 22	0.0278 38	0.0672 40	-0.0696 73	0.0297 50	-0.0035 57
O(3)	0.0249 20	0.0187 32	0.0340 26	-0.0127 59	0.0123 39	0.0107 50
O(4)	0.0378 25	0.0263 38	0.0367 28	-0.0106 66	0.0269 45	0.0180 61
C(1)	0.0171 26	0.0342 62	0.0280 35	0.0054 90	0.0103 51	-0.0179 71
C(2)	0.0310 35	0.0531 79	0.0343 41	-0.0355 107	0.0263 63	0.0165 98
C(3)	0.0323 36	0.0493 83	0.0483 50	-0.0288 121	0.0301 73	0.0090 101
C(4)	0.0430 45	0.0583 95	0.0826 81	-0.0486 172	0.0526 103	0.0453 128
C(5)	0.0244 29	0.0349 59	0.0258 34	-0.0327 87	0.0228 53	0.0135 83
C(6)	0.0249 31	0.0370 62	0.0416 42	-0.0222 100	0.0344 61	0.0178 85
C(7)	0.0166 26	0.0365 64	0.0402 41	-0.0317 98	0.0218 55	0.0150 76
C(8)	0.0473 47	0.0646 100	0.0694 66	-0.0148 156	0.0757 95	0.0787 129

TABLE 5.7

PRINCIPAL VALUES AND DIRECTIONS OF
VIBRATION TENSORS
WITH RESPECT TO ORTHOGONAL AXES

ATOM	^{o2} U A	D1	D2	D3
K(1)	0.0415	0.6662	-0.0946	0.7397
	0.0137	-0.5290	0.6392	0.5582
	0.0262	-0.5257	-0.7632	0.3758
O(1)	0.0623	-0.1301	-0.4972	0.8578
	0.0344	0.9601	0.1529	0.2342
	0.0470	-0.2476	0.8541	0.4575
O(2)	0.0688	0.2032	-0.5353	0.8199
	0.0091	-0.2387	0.7850	0.5717
	0.0254	-0.9496	-0.3119	0.0317
O(3)	0.0154	-0.5172	0.8515	0.0863
	0.0377	0.2643	0.0629	0.9624
	0.0274	0.8141	0.5205	-0.2576
O(4)	0.0512	0.6947	0.3415	0.6331
	0.0211	-0.5353	0.8334	0.1377
	0.0306	-0.4806	-0.4345	0.7617
C(1)	0.0610	-0.1660	0.6241	0.7635
	0.0115	-0.7744	-0.5619	0.2909
	0.0228	0.6105	-0.5430	0.5765
C(2)	0.0183	-0.6922	0.1403	0.7080
	0.0568	0.3709	0.9106	0.1821
	0.0376	-0.6192	0.3886	-0.6823
C(3)	0.0266	-0.8977	0.1017	0.4288
	0.0580	0.4061	0.5686	0.7154
	0.0455	-0.1711	0.8163	-0.5517
C(4)	0.0136	-0.7771	0.4637	0.4257
	0.0924	0.5626	0.2083	0.8001
	0.0695	0.2823	0.8612	-0.4227

c(5)	0.0579 0.0383 0.0299	-0.5705 0.4441 -0.6909	0.2784 0.8960 0.3461	0.7727 0.0051 -0.6348
c(6)	0.0116 0.0528 0.0361	-0.8223 0.5689 -0.0160	0.2587 0.3986 0.8799	0.5069 0.7194 -0.4750
c(7)	0.0087 0.0406 0.0391	-0.8648 0.4244 0.2684	0.2879 -0.0188 0.9575	0.4114 0.9053 -0.1059
c(8)	0.1251 -0.0523 0.0549	0.6236 -0.7814 0.0253	0.5115 0.3833 -0.7691	0.5912 0.4925 0.6387

Tables of Bond-Lengths, Bond-Angles and Interatomic

Contacts

In the following tables, coordinates of atoms (x, y, z) denoted only by a bracketed number are those given in TABLE 5.1. The others are related to this set as follows:

$$A - l + x, y, z$$

$$B - x, y - l, z$$

$$C - x - l, y, z$$

$$D - x, y + l, z$$

$$E - x - l, y - l, z$$

TABLE 5.8

INTERATOMIC DISTANCES IN ANGSTROMS

O(1) - C(1)	1.242 ± 13
C(1) - O(2)	1.321 ± 10
C(1) - C(2)	1.478 ± 11
C(2) - C(3)	1.267 ± 16
C(3) - C(4)	1.447 ± 14
O(3) - C(5)	1.308 ± 10
C(5) - O(4)	1.277 ± 13
C(5) - C(6)	1.450 ± 11
C(6) - C(7)	1.320 ± 16
C(7) - C(8)	1.483 ± 13
O(2)O(3A)	2.462 ± 10
K(1)O(4B)	2.640 ± 6
K(1)O(1C)	2.690 ± 7
K(1)O(3)	2.786 ± 6
K(1)O(4D)	2.802 ± 6
K(1)O(2E)	2.816 ± 6

TABLE 5.9

BOND-ANGLES AND E.S.D.

O(1) - C(1) - O(2)	125.3 ± 7
O(1) - C(1) - C(2)	118.8 ± 7
O(2) - C(1) - C(2)	115.9 ± 8
C(1) - C(2) - C(3)	123.9 ± 8
C(2) - C(3) - C(4)	122.5 ± 9
O(3) - C(5) - C(6)	118.9 ± 9
O(4) - C(5) - C(6)	116.8 ± 8
C(5) - C(6) - C(7)	123.7 ± 9
C(6) - C(7) - C(8)	125.8 ± 9

TABLE 5.10

Final observed and calculated structure factors.

FIG. 3

Projection of the structure, from $z = -0.5$ to $z = 1.0$, onto the xy plane. This projection is perpendicular to the plane, i.e. down the c^* axis. Molecules represented by double circles have the largest z -coordinates, those represented by dotted circles have the smallest z -coordinates.

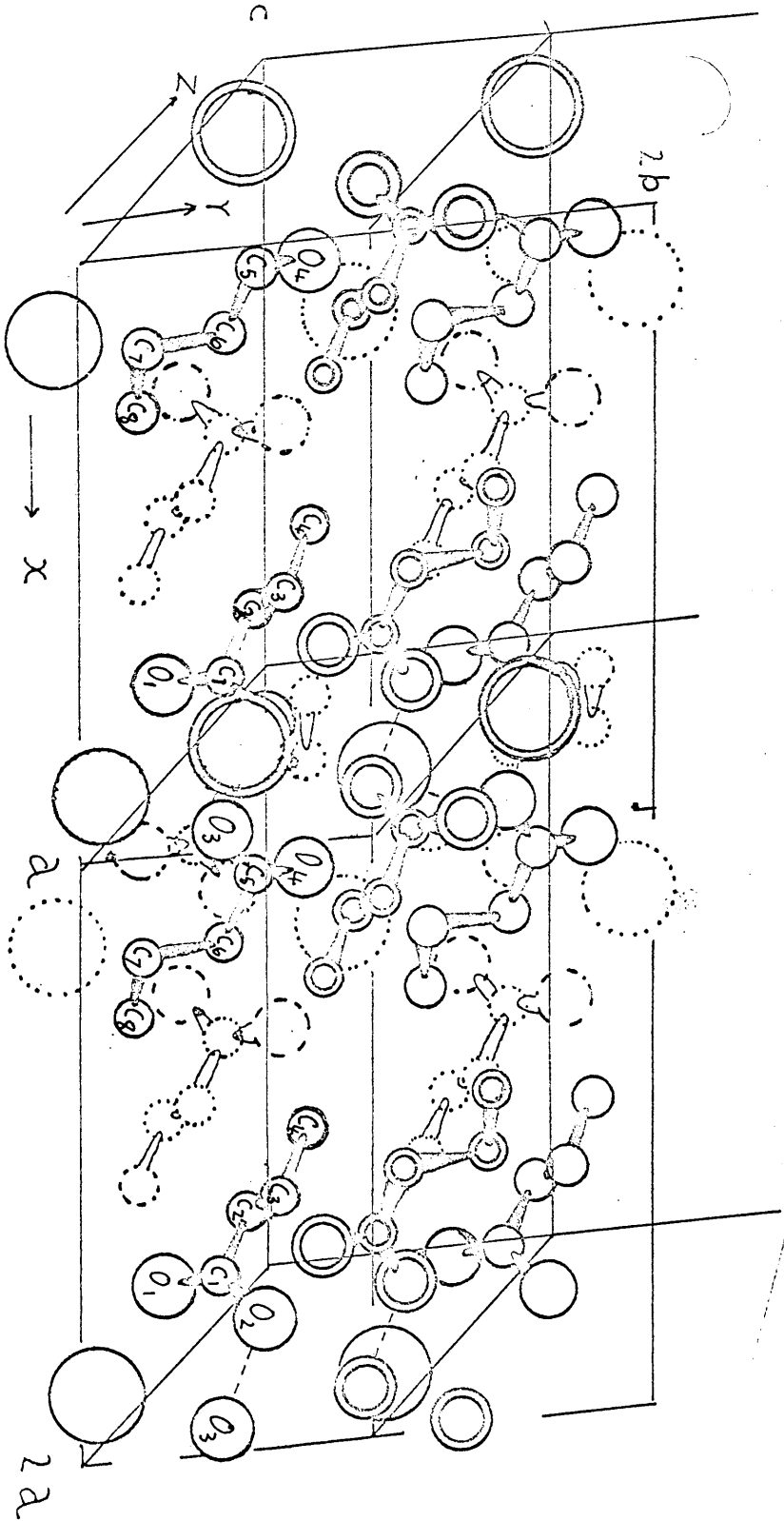
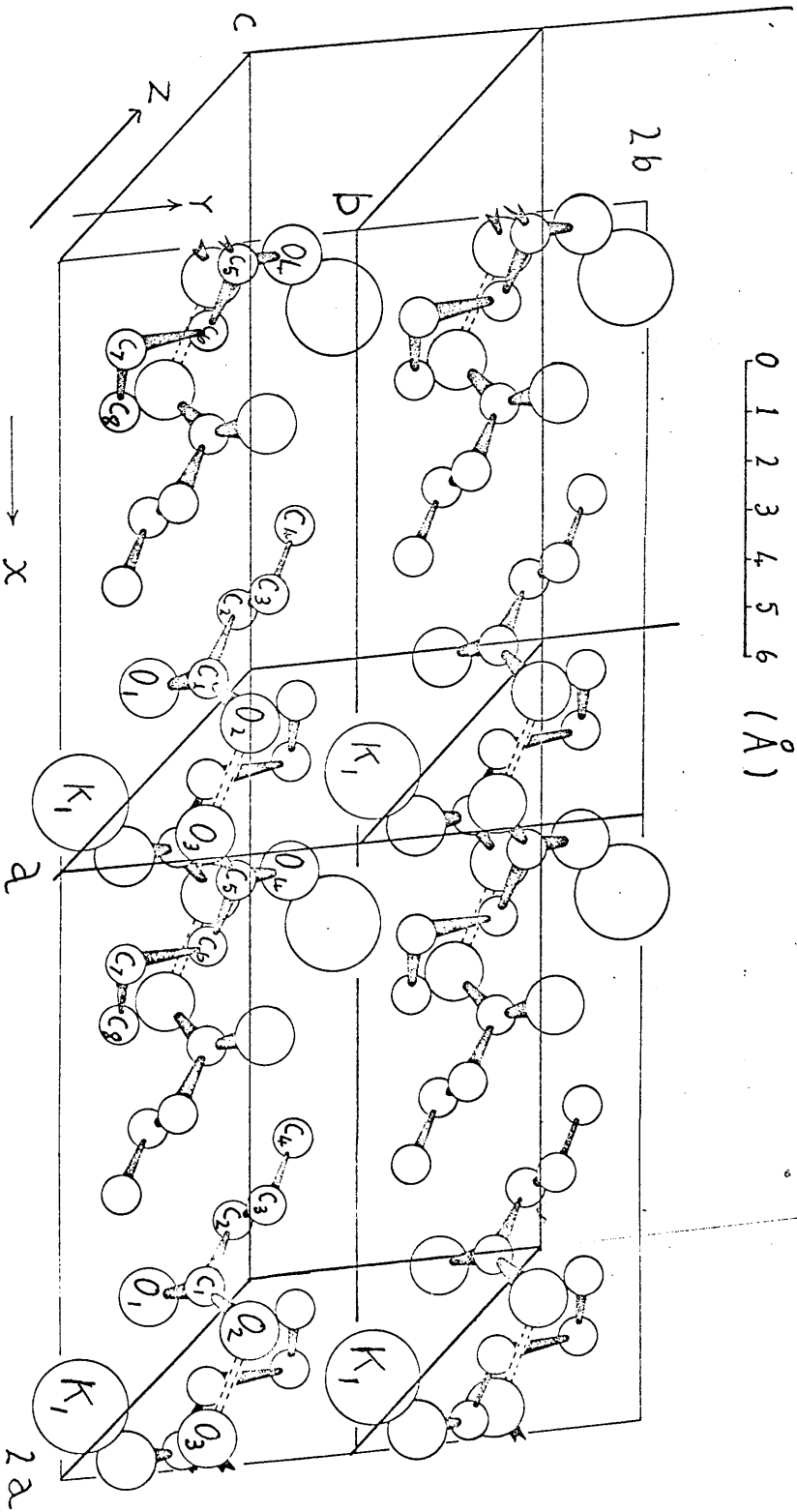


FIG. 4

Projection of the structure from $z = -0.5$ to $z = 0.5$
onto the xy plane, down the c^* axis.



0 1 2 3 4 5 6 (\AA)

TABLE 5.11

MEAN PLANES

Plane 1

$$O(1) - O(2) - C(1) - C(2)$$

$$0.0035 X + 0.6062 Y - 0.7953 Z = 0.8386 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Root Mean Square Deviation from Plane} = 0.006 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Maximum Deviation (C(1))} = 0.011 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

Plane 2

$$C(1) - C(2) - C(3) - C(4)$$

$$0.0382 X + 0.6423 Y - 0.7655 Z = 1.230 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Root Mean Square Deviation from Plane} = 0.00018 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{All Deviations} = 0.0002 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

Plane 3

$$O(3) - O(4) - C(5) - C(6)$$

$$-0.4911 X - 0.1442 Y - 0.8591 Z = -3.1467 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Root Mean Square Deviation from Plane} = 0.0041 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Maximum Deviation (C(5))} = 0.007 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

Plane 4

$$C(5) - C(6) - C(7) - C(8)$$

$$-0.5008 X - 0.0754 Y - 0.8623 Z = -3.0068 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Root Mean Square Deviation from Plane} = 0.018 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

$$\text{Maximum Deviation (C(7))} = 0.019 \text{ (}\overset{\circ}{\text{A}}\text{)}$$

ANGLES

$$\text{Plane 1 - Plane 2} = 3.34^\circ$$

$$\text{Plane 3 - Plane 4} = 3.98^\circ$$

not $P1$, an attempt was made to refine a structure, equivalent to the final structure, in the latter space-group. After two cycles of Least-Squares Refinement the R-factor had dropped from 15.2% to 15.08%. No significant change took place in any of the atomic coordinate parameters, but their standard deviations became much larger because the ratio of parameters to observations was greatly increased. It was therefore concluded that the space group is indeed $P\bar{1}$.

5.6. Discussion

The more important bond-lengths, bond-angles and interatomic distances are shown on Fig. 1.

This structure has a very short hydrogen bond ($2.46 \pm .01\text{\AA}$). Bonds of this length are most often found in symmetrical acid salts of Type A (cf. potassium hydrogen dianisate). Hydrogen bonds joining crystallographically distinguishable acid residues are usually considerably longer (2.53\AA).

From the molecular packing diagram it can be seen that there does, in fact, appear to be a pseudo-centre of symmetry situated on the mid-point of the hydrogen bond between the oxygen atoms. From the present X-ray structure analysis the accuracy is, unfortunately, not high enough to distinguish between the anion-like residue and the acid-like residue. Nevertheless it can be seen that the residues are much more similar than most "unsymmetrical" Type B acid salts.

As in the case of ammonium hydrogen dicinnamate (Bryan, Mills and Speakman, 1963) some subtle form of disorder may be present in this structure. Crotonic and cinnamic acid have similar unsymmetrical "tails" attached to the carboxyl group, and these may possibly prevent efficient packing of the molecules. In all three compounds,

ammonium hydrogen dicinnamate, potassium hydrogen
dicrotonate and rubidium hydrogen dicrotonate, disorder
of the "tails" has been suspected.

5.7. Appendix

Reduced cell:

$$a = 12.92 \text{ \AA}, \quad b = 6.02 \text{ \AA}, \quad c = 7.48 \text{ \AA}, \\ \alpha = 114^\circ 5', \quad \beta = 103^\circ 29', \quad \gamma = 98^\circ 11'.$$

Cell used in this analysis:

$$a = 12.46 \text{ \AA}, \quad b = 6.02 \text{ \AA}, \quad c = 7.45 \text{ \AA}, \\ \alpha = 66^\circ 31', \quad \beta = 103^\circ 39', \quad \gamma = 95^\circ 12'.$$

Transformation matrix for transforming the arbitrary cell to the reduced cell:

$$\begin{bmatrix} 1 & 0 & \bar{1} \\ 0 & \bar{1} & 0 \\ 0 & 1 & \bar{1} \end{bmatrix}$$

THE CRYSTAL STRUCTURE DETERMINATION OF RUBIDIUM
HYDROGEN DICROTONATE

Contents

- 6.1. Introduction
- 6.2. Experimental
- 6.3. Structure Solution
- 6.4. Discussion

Chapter 6

THE CRYSTAL STRUCTURE DETERMINATION

OF

RUBIDIUM HYDROGEN DICROTONATE

6.1. Introduction

The complex crystal structure of this chemically simple acid salt has been the main feature of interest throughout the structure determination by X-ray diffraction. This structure determination is not of high accuracy, mainly because of the poor quality of the crystals, the rapid fall-off of the X-ray diffracted intensities with $\sin \theta$ and the large ratio of parameters to observations. There are four rubidium atoms and eight associated acid residues in the asymmetric portion of the unit cell. These 52 atoms (excluding hydrogen) had to be positioned, making a total of 236 parameters (including thermal-vibration-parameters and layer-scale-factors) to be determined. 1269 independent reflexions were measured, thus giving a ratio of parameters:observations of approximately 1:5.4. It is suspected that a certain amount of disorder occurs in the positions occupied by the hydrocarbon "tails" of the molecules, and this, combined with the relatively small number of observations, and the large number of atomic

parameters to be determined, has resulted in the low accuracy of the present determination. Nevertheless, it is considered that the over-all structure is certainly correct, and is the most complex acid-salt structure which has so far been determined.

6.2. Experimental

The crystals were of poor quality. They had a soft texture, exhibited poor extinction between crossed Nicol prisms, and were so deliquescent that small fragments would often dissolve while being examined on the microscope stage. Selected crystals were mounted individually in thin-walled capillary glass tubes, which were then sealed at both ends.

Some difficulty was experienced in selecting a small single crystal. About a dozen different crystals were mounted, and set, only to be discarded when X-ray photographs revealed that the crystal was split, or that several "satellite" crystals were adhering to the main one. The crystal finally selected, although not perfect, had only a single small "satellite".

Crystal Data

Weissenberg photographs were taken about both the shorter axes. The following parameters were obtained from photographs taken with Cu K_{α} radiation:

Rubidium hydrogen dicrotonate $\text{Rb H}(\text{C}_4\text{H}_5\text{O}_2)_2$; $M = 259$;
orthorhombic; $a = 33.10$; $b = 17.46$; $c = 7.62\text{\AA}$;
 $V = 4404\text{\AA}^3$; $D_m = 1.52$; $z = 16$; $D_c = 1.52$; absorption
co-efficient for (Cu) X-rays $\mu = 63\text{cm}^{-1}$.

Choice of Space Groups

This was made more difficult by the presence of "Systematic weaknesses" which, had they been true, would have indicated the space group Pccn. A careful examination of the photographs enabled the crystal to be assigned to the space group $P2_12_12$ which is a subgroup of Pccn.

6.3. Structure Solution

Two two-dimensional Patterson maps were calculated for the projections (u, v, o) and (u, o, w). Both were expected to have the symmetry of the plane group pmm. The plane group of the (u, v, o) projection, however, was found to approximate to cmm, while the (u, o, w) projection had a pseudo-mirror line at $w = \frac{1}{4}$. From the pseudo-symmetry, and the systematically weak reflexions noted above, it was inferred that the heavy-atoms, at least, were in a more symmetrical arrangement than that required by the true space group.

"Sharpened" and "unsharpened" three-dimensional Patterson maps were then calculated, and the positions of four rubidium atoms were obtained, which accounted for all the major peaks of the maps. The initial structure-factors, calculated for the heavy-atoms alone, gave an R-factor of 48%.

The electron-density map calculated from the observed structure-amplitudes with the heavy-atom phases, had pseudo-symmetry (pseudo centre-of-symmetry) and only two crotonate residues could be recognised. The structure solution proceeded slowly. After seven structure-factor- $|F_o|$ -electron-density map cycles, a "Bunn Error Synthesis," and one $|F_c|$ electron-density map, only 32 atoms had

been located, and the R-factor had only dropped to 30%. Although the positions of the unknown atoms were indicated by complex regions of electron-density, it was impossible to decide how the actual molecules were positioned. The $|F_o|$ map showed that part of the trouble was caused by diffraction-ripples from the heavy-atoms. Another cause of difficulty was the continued pseudo-symmetry of the structure.

It was therefore decided to refine the partial structure by the Least-Squares Method. R. Pollard's (1966) Least-Squares Program was used, with individual isotropic atomic temperature parameters, individual batch scale factors for each reciprocal lattice net, and unit weights for every observation. After four cycles this refinement of the partial structure converged to an R-factor of 22%, but by now the atomic co-ordinates had been shifted considerably by the refinement.

An $(F_o - F_c)$ map was now calculated. Its most easily interpretable features were large characteristic "clover-leaf" patterns at the site of each heavy atom.

It was therefore decided to continue the Least-Squares Refinement of the partial structure, with anisotropic temperature parameters for the heavy atoms.

Three cycles of refinement with Cruickshank's (1965)

Least-Squares Program then reduced the R-factor to 18%. The Least-Squares Refinement destroyed the pseudo-symmetry, and the observed and calculated structure amplitudes were now in much better agreement. The severe rippling remained. Thus the problem was to differentiate the unknown atoms of the structure from the diffraction ripples.

At this stage, the usual course would be to calculate an $(F_o - F_c)$ map. If the atoms of the known acid residues and carboxyl groups are included in the phasing calculation, they do not appear on the difference map, and it becomes difficult to link-up the isolated peaks which do appear. If, however, the light atoms are excluded from the phasing calculation, the pseudo-symmetry of the heavy-atom arrangement causes confusing pseudo-symmetry in the resulting difference map.

It was decided that what was required was a map, effectively an electron-density map of the light-atom regions of the unit cell, and a difference map of the heavy-atom structure. Thus the heavy-atom diffraction ripples would also be removed from the map. A map was calculated, using Fourier coefficients AD and BD such that:

$$AD = A_o - AH,$$

and

$$BD = B_o - BH,$$

$$A_o = \frac{A_c \times |F_o|}{|F_c|},$$

$$B_o = \frac{B_c \times |F_o|}{|F_c|},$$

where A_c and B_c , AH and BH are the real and imaginary components of the calculated structure factors for all the known atoms, and only the heavy-atoms, respectively. $|F_o|$ and $|F_c|$ are the observed and calculated structure amplitudes.

The map degenerates to a difference-map if the positions of only the heavy-atoms are known. The modified co-efficients AD and BD were calculated by a small Algol program. The map produced from these co-efficients was very clean in appearance. The light-atom structure showed up clearly, and it was possible to locate the remaining atoms, bringing the total number of atoms in the asymmetric portion of the unit cell to fifty-two.

The whole structure was then given two cycles of Least-Squares refinement, as a result of which the R-factor fell to 16.5%. Structure-factors were then calculated, using R. Pollard's (1966) program, and a weighting scheme:

$$w = K \exp -\{A + B F_o + C F_o^2 + D \lambda^2 / \sin^2 \theta + E \lambda^4 / \sin^4 \theta + F \sin^2 \theta / \lambda^2\},$$

was fitted by the program described in Chapter 2.

Least-Squares Refinement was then continued for a further three cycles. The parameter standard deviations fell steadily, although they were still large compared with those normally encountered at this stage of a structure analysis.

Anisotropic structure-factors were again calculated, and a final ($F_o - F_c$) map was obtained. This showed no major features, although the background fluctuated in the region $\pm 0.8 \text{ \AA}^3$, and a complex region surrounded each heavy-atom site. Since, during the course of the refinement, the distribution of $W \Delta^2$ had again become uneven, the weighting scheme program was used to re-adjust the parameters of the weighting scheme. Two further cycles of least-squares concluded the refinement. The final R-factor was 13.5%.

6.4. Discussion

The atomic parameters are shown in Tables (6.1) to (6.5); intramolecular bond lengths with their estimated standard deviations are shown in Table (6.6) and rubidium oxygen and oxygen oxygen contacts are given in Table (6.8). Average bond-lengths, (assuming all eight acid residues in the asymmetric portion of the unit cell to be chemically equivalent) are given in Table (6.10) along with their root-mean-square deviations and average standard deviations calculated from the Least-Squares Estimated Standard Deviations of the atomic parameters.

This crystal structure, space group $P2_1 2_1 2$ approximates to the more symmetrical space group Pccn, which can be derived from $P2_1 2_1 2$ by placing a centre of symmetry at $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$. To obtain the conventional representation of Pccn the origin of the crystal axes is then moved to the centre of symmetry.

If the origin of the axes is transferred to $(\frac{1}{2}, \frac{1}{2}, \frac{1}{2})$ in the present structure, it can be seen from the resulting fractional atomic co-ordinates how closely the structure's symmetry approximates to Pccn. If the space group was, in fact, Pccn, there would be only two rubidium atoms and four acid residues in the asymmetric

portion of the unit cell, and these would be related by the centre of symmetry to two other rubidium atoms and to another four acid residues.

The majority of acid salts with very short hydrogen bonds have been found to have a symmetrical structure in which the two acid residues linked by the hydrogen bond are crystallographically equivalent. The bond lies across a centre of symmetry, or a diad axis, and the acid residues cannot be distinguished as "acid molecule" and "anion". These are known as "Class A" acid salts. Where the two residues linked by the hydrogen bond are not themselves symmetry-related, the hydrogen bond is usually longer than in "symmetrical" acid salts and the residues can be distinguished as "acid molecule" and "anion" respectively. These are known as "Class B" acid salts.

In rubidium hydrogen dicrotonate, the hydrogen bonds, though short, do not connect crystallographically equivalent residues, nor do they lie across the pseudo-centre of-symmetry of the pseudo Pccn space group. The criterion for a residue being an anion is:-

Symmetry of the carboxyl group (equal C-O bonds and equal C-C-O angles).

The criteria for recognising a residue as an acid

molecule are:

1. Distinguishable unequal C = O and C - O ... H bonds.
2. The bond-angles in the carboxyl groups are unequal
 $C - C - O \dots H < C - C = O$.

Unfortunately this structure determination is not sufficiently accurate to observe the distinction. It must be noted, however, that while in not one of the eight acid residues are the carbon-oxygen bonds of the carboxyl group equal, in every case the hydrogen bond is attached through the oxygen with the longer bond. In addition the average O ... O distance for the hydrogen bonded atoms is $2.47 \pm .06 \text{ \AA}$ which compares with 2.462 \AA in potassium hydrogen dicrotonate where, in spite of greater accuracy than in the present determination, again no distinction can be made between molecule and anion.

Thus, although these bonds are as short as many hydrogen bonds found in symmetrical "Class A" structures, it is interesting to observe that in neither the potassium hydrogen dicrotonate nor the rubidium hydrogen dicrotonate have the molecules packed so that the hydrogen bond could occupy a symmetrical environment, although in the potassium compound the space group is $P \bar{1}$, with a centre of symmetry available, and in the rubidium compound the pseudo-centre of symmetry is ignored.

Thus rubidium hydrogen dicrotonate and potassium hydrogen dicrotonate, although simple chemical compounds, have entirely different crystal structures and pack in different ways. This contrasts with the behaviour of many organic acids which form series of isomorphous acid salts with different cations. This change of structure, and the extremely deliquescent behaviour of the rubidium compound, indicate that neither of these structures has good packing stability. This must be due to the nature of the carbon chain attached to the carboxyl group. It is less symmetrical than any of the derivatives of benzoic acid or acetic acid.

It is interesting to note that the structure analysis of ammonium hydrogen dicinnamate ($(\text{Ph.CH.CH.CO}^2)_2\text{H NH}_4$, Bryan, Mills and Speakman, 1963), was hampered by the rapid fall-off of intensity with $\sin\theta$, and that the corresponding molecular dimensions are anomalous. Here, too, the authors suspected "some subtle disorder" of the carbon chain.

The cinnamic acid residue has exactly the same "kink" in the "tail" as the crotonate residue. It is felt that in the rubidium hydrogen dicrotonate structure disorder may be present, particularly with regard to the outermost carbon atoms of the acid residue "tails". As can be seen from Table 6.3, these all have large vibration parameters,

and these large values may be caused by a partial disorder in the "tail" portions.

The co-ordination of the rubidium ions is also anomalous. The contacts are listed in Table 6.8.

TABLE 6.1

FRACTIONAL COORDINATES AND E.S.D.

ATOM	X	Y	Z
Rb(1)	0.2900 \pm 2	0.4245 \pm 3	0.8210 \pm 9
Rb(2)	0.3197 \pm 2	0.3258 \pm 3	0.3203 \pm 10
Rb(3)	0.2135 \pm 2	0.0615 \pm 3	0.6800 \pm 9
Rb(4)	0.1806 \pm 2	0.1609 \pm 3	0.1836 \pm 10
O(1)	0.1586 \pm 11	0.0351 \pm 21	0.9874 \pm 49
O(2)	0.1390 \pm 13	-0.0854 \pm 23	0.9758 \pm 57
C(1)	0.1297 \pm 13	-0.0070 \pm 26	0.9646 \pm 66
C(2)	0.0902 \pm 20	0.0151 \pm 39	0.9389 \pm 96
C(3)	0.0734 \pm 26	0.0890 \pm 46	0.9455 \pm 115
C(4)	0.0289 \pm 31	0.1136 \pm 54	0.9070 \pm 138
O(3)	0.2785 \pm 12	0.2736 \pm 21	0.9951 \pm 52
O(4)	0.2405 \pm 16	0.1656 \pm 28	0.9255 \pm 70
C(5)	0.2777 \pm 23	0.1971 \pm 42	0.9170 \pm 103
C(6)	0.3156 \pm 30	0.1788 \pm 51	0.8049 \pm 150
C(7)	0.3476 \pm 25	0.2258 \pm 44	0.8355 \pm 122
C(8)	0.3776 \pm 27	0.2067 \pm 50	0.6602 \pm 135
O(5)	0.2458 \pm 13	0.3251 \pm 23	0.5690 \pm 58
O(6)	0.2260 \pm 12	0.2102 \pm 22	0.5059 \pm 55
C(9)	0.2204 \pm 16	0.2737 \pm 30	0.5938 \pm 71
C(10)	0.1840 \pm 28	0.3071 \pm 49	0.6413 \pm 131

C(11)	0.1552 ± 22	0.2652 ± 39	0.7235 ± 96
C(12)	0.1138 ± 24	0.2918 ± 43	0.7702 ± 104
O(7)	0.2489 ± 13	0.4226 ± 24	0.1521 ± 64
O(8)	0.2138 ± 18	0.3155 ± 32	0.1200 ± 78
C(13)	0.2133 ± 26	0.4029 ± 43	0.1931 ± 120
C(14)	0.1820 ± 27	0.4422 ± 47	0.3307 ± 135
C(15)	0.1467 ± 27	0.4049 ± 48	0.3258 ± 133
C(16)	0.1185 ± 35	0.4501 ± 68	0.4167 ± 172
O(9)	0.2850 ± 12	0.1627 ± 21	0.3651 ± 51
O(10)	0.2524 ± 13	0.0640 ± 24	0.2991 ± 65
C(17)	0.2808 ± 20	0.0994 ± 35	0.2750 ± 82
C(18)	0.3214 ± 25	0.0581 ± 43	0.2299 ± 111
C(19)	0.3487 ± 26	0.0895 ± 47	0.2002 ± 131
C(20)	0.3881 ± 26	0.0402 ± 52	0.1014 ± 126
O(11)	0.3702 ± 13	0.5812 ± 25	0.5633 ± 64
O(12)	0.3397 ± 17	0.4671 ± 32	0.5440 ± 80
C(21)	0.3714 ± 24	0.5064 ± 46	0.5684 ± 117
C(22)	0.4109 ± 27	0.4812 ± 52	0.5947 ± 124
C(23)	0.4216 ± 25	0.4132 ± 44	0.6316 ± 108
C(24)	0.4665 ± 26	0.4127 ± 46	0.6764 ± 136
O(13)	0.0955 ± 13	0.2249 ± 24	0.2440 ± 55
O(14)	0.0736 ± 14	0.1690 ± 25	0.4183 ± 62
C(25)	0.0671 ± 26	0.2200 ± 46	0.3329 ± 131
C(26)	0.0322 ± 25	0.2788 ± 47	0.3371 ± 133

C(27)	0.0412 ± 35	0.3378 ± 57	0.2028 ±168
C(28)	0.0003 ± 43	0.3615 ± 64	0.2919 ±172
O(15)	0.3910 ± 14	0.2581 ± 25	0.2588 ± 57
O(16)	0.4240 ± 15	0.3246 ± 27	0.0023 ± 66
C(29)	0.4204 ± 21	0.2679 ± 40	0.1476 ±101
C(30)	0.4618 ± 24	0.2267 ± 43	0.1274 ±105
C(31)	0.4599 ± 32	0.1721 ± 54	0.2548 ±136
C(32)	0.5022 ± 35	0.1194 ± 56	0.3116 ±151

TABLE 6.2

ATOMIC COORDINATES AND E.S.D.
IN ANGSTROMS

ATOM	X	Y	Z
Rb(1)	9.60 ± 1	7.41 ± 1	6.26 ± 1
Rb(2)	10.58 ± 1	5.69 ± 1	2.44 ± 1
Rb(3)	7.07 ± 1	1.07 ± 0	5.18 ± 1
Rb(4)	5.98 ± 1	2.81 ± 1	1.40 ± 1
O(1)	5.25 ± 4	0.61 ± 4	7.52 ± 4
O(2)	4.60 ± 4	-1.49 ± 4	7.44 ± 4
C(1)	4.29 ± 4	-0.12 ± 5	7.35 ± 5
C(2)	2.98 ± 7	0.26 ± 7	7.15 ± 7
C(3)	2.43 ± 9	1.55 ± 8	7.20 ± 9
C(4)	0.96 ± 10	1.98 ± 9	6.91 ± 11
O(3)	9.22 ± 4	4.78 ± 4	7.58 ± 4
O(4)	7.96 ± 5	2.89 ± 5	7.05 ± 5
C(5)	9.19 ± 8	3.44 ± 7	6.99 ± 8
C(6)	10.45 ± 10	3.12 ± 9	6.13 ± 11
C(7)	11.50 ± 8	3.94 ± 8	6.37 ± 9
C(8)	12.50 ± 9	3.61 ± 9	5.03 ± 10
O(5)	8.14 ± 4	5.68 ± 4	4.34 ± 4
O(6)	7.48 ± 4	3.67 ± 4	3.86 ± 4
C(9)	7.30 ± 5	4.78 ± 5	4.52 ± 5
C(10)	6.09 ± 9	5.36 ± 9	4.89 ± 10

C(11)	5.14	±	7	4.63	±	7	5.51	±	7
C(12)	3.77	±	8	5.09	±	8	5.87	±	8
O(7)	8.24	±	4	7.38	±	4	1.16	±	5
O(8)	7.08	±	6	5.51	±	6	0.91	±	6
C(13)	7.06	±	9	7.04	±	8	1.47	±	9
C(14)	6.02	±	9	7.72	±	8	2.52	±	10
C(15)	4.85	±	9	7.07	±	8	2.48	±	10
C(16)	3.92	±	12	7.86	±	12	3.18	±	13
O(9)	9.43	±	4	2.84	±	4	2.78	±	4
O(10)	8.36	±	4	1.12	±	4	2.28	±	5
C(17)	9.30	±	7	1.74	±	6	2.10	±	6
C(18)	10.64	±	8	1.02	±	8	1.75	±	8
C(19)	11.54	±	9	1.56	±	8	1.53	±	10
C(20)	12.85	±	9	0.70	±	9	0.77	±	10
O(11)	12.25	±	4	10.15	±	4	4.29	±	5
O(12)	11.24	±	6	8.16	±	6	4.15	±	6
C(21)	12.29	±	8	8.84	±	8	4.33	±	9
C(22)	13.60	±	9	8.40	±	9	4.53	±	9
C(23)	13.96	±	8	7.21	±	8	4.81	±	8
C(24)	15.44	±	9	7.21	±	8	5.15	±	10
O(13)	3.16	±	4	3.93	±	4	1.86	±	4
O(14)	2.44	±	5	2.95	±	4	3.19	±	5
C(25)	2.22	±	9	3.84	±	8	2.54	±	10
C(26)	1.07	±	8	4.87	±	8	2.57	±	10

C(27)	1.36 ± 12	5.90 ± 10	1.55 ± 13
C(28)	0.01 ± 14	6.31 ± 11	2.22 ± 13
O(15)	12.94 ± 5	4.51 ± 4	1.97 ± 4
O(16)	14.03 ± 5	5.67 ± 5	0.02 ± 5
C(29)	13.92 ± 7	4.68 ± 7	1.12 ± 8
C(30)	15.28 ± 8	3.96 ± 7	0.97 ± 8
C(31)	15.22 ± 11	3.00 ± 9	1.94 ± 10
C(32)	16.62 ± 12	2.08 ± 10	2.37 ± 12

TABLE 6.3FRACTIONAL COORDINATES
AND TEMPERATURE FACTORS
FOR CARBON AND OXYGEN ATOMS

ATOM	X	Y	Z	U
O(1)	0.1586	0.0351	0.9874	0.0508
O(2)	0.1390	-0.0854	0.9758	0.0670
C(1)	0.1297	-0.0070	0.9646	0.0250
C(2)	0.0902	0.0151	0.9389	0.0732
C(3)	0.0734	0.0890	0.9455	0.0962
C(4)	0.0289	0.1136	0.9070	0.1212
O(3)	0.2785	0.2736	0.9951	0.0607
O(4)	0.2405	0.1656	0.9255	0.0967
C(5)	0.2777	0.1971	0.9170	0.0818
C(6)	0.3156	0.1788	0.8049	0.1221
C(7)	0.3476	0.2258	0.8355	0.0951
C(8)	0.3776	0.2067	0.6602	0.1094
O(5)	0.2458	0.3251	0.5690	0.0699
O(6)	0.2260	0.2102	0.5059	0.0648
C(9)	0.2204	0.2737	0.5938	0.0422
C(10)	0.1840	0.3071	0.6413	0.1084
C(11)	0.1552	0.2652	0.7235	0.0734
C(12)	0.1138	0.2918	0.7702	0.0841
O(7)	0.2489	0.4226	0.1521	0.0775

o(8)	0.2138	0.3155	0.1200	0.1133
c(13)	0.2133	0.4029	0.1931	0.0974
c(14)	0.1820	0.4422	0.3307	0.1099
c(15)	0.1467	0.4049	0.3258	0.1087
c(16)	0.1185	0.4501	0.4167	0.1311
o(9)	0.2850	0.1627	0.3651	0.0606
o(10)	0.2524	0.0640	0.2991	0.0822
c(17)	0.2808	0.0994	0.2750	0.0618
c(18)	0.3214	0.0581	0.2299	0.0915
c(19)	0.3487	0.0895	0.2002	0.1050
c(20)	0.3881	0.0402	0.1014	0.1082
o(11)	0.3702	0.5812	0.5633	0.0797
o(12)	0.3397	0.4671	0.5440	0.1119
c(21)	0.3714	0.5064	0.5684	0.0935
c(22)	0.4109	0.4812	0.5947	0.1101
c(23)	0.4216	0.4132	0.6316	0.0913
c(24)	0.4665	0.4127	0.6764	0.1046
o(13)	0.0955	0.2249	0.2440	0.0705
o(14)	0.0736	0.1690	0.4183	0.0782
c(25)	0.0671	0.2200	0.3329	0.1040
c(26)	0.0322	0.2788	0.3371	0.1009
c(27)	0.0412	0.3378	0.2028	0.1321
c(28)	0.0003	0.3615	0.2919	0.1297
o(15)	0.3910	0.2581	0.2588	0.0745

0(16)	0.4240	0.3246	0.0023	0.0835
C(29)	0.4204	0.2679	0.1476	0.0787
C(30)	0.4618	0.2267	0.1274	0.0839
C(31)	0.4599	0.1721	0.2548	0.1224
C(32)	0.5022	0.1194	0.3116	0.1338

TABLE 6.4

ANISOTROPIC TEMPERATURE FACTORS AND E.S.D.
FOR RUBIDIUM ATOMS

ATOM	U11	U22	U33	2U23	2U31	2U12
Rb(1)	0.1066 49	0.0427 31	0.0474 39	0.0059 63	-0.0094 90	-0.0287 74
Rb(2)	0.0786 40	0.0443 33	0.0702 45	0.0009 75	-0.0020 87	-0.0272 69
Rb(3)	0.0864 40	0.0357 30	0.0565 40	0.0102 64	0.0068 87	-0.0135 65
Rb(4)	0.0712 38	0.0498 33	0.0700 44	-0.0059 76	-0.0219 86	-0.0311 70

TABLE 6.5

PRINCIPAL VALUES AND DIRECTIONS OF
VIBRATION TENSORS

ATOM	^{o2}	D1	D2	D3
	U A			
Rb(1)	0.1101	-0.9740	0.2110	0.0829
	0.0392	-0.1866	-0.9539	0.2352
	0.0474	0.1287	0.2136	0.9684
Rb(2)	0.0396	-0.3289	-0.9444	0.0031
	0.0834	-0.9411	0.3280	0.0823
	0.0701	-0.0787	0.0241	-0.9966
Rb(3)	0.0875	0.9838	-0.1201	0.0886
	0.0335	-0.1378	-0.9625	0.2337
	0.0576	0.0572	-0.2433	-0.9683
Rb(4)	0.0396	0.4967	0.8284	0.2589
	0.0846	-0.7901	0.3080	0.5300
	0.0668	0.3593	-0.4678	0.8075

TABLE 6.6

BOND-LENGTHS WITH E.S.D.

O(1) - C(1)	1.22	±	6
O(2) - C(1)	1.41	±	6
C(1) - C(2)	1.38	±	8
C(2) - C(3)	1.40	±	11
C(3) - C(4)	1.56	±	13
O(3) - C(5)	1.46	±	8
O(4) - C(5)	1.35	±	9
C(5) - C(6)	1.55	±	13
C(6) - C(7)	1.36	±	12
C(7) - C(8)	1.69	±	13
O(5) - C(9)	1.24	±	7
O(6) - C(9)	1.31	±	7
C(9) - C(10)	1.39	±	11
C(10) - C(11)	1.35	±	12
C(11) - C(12)	1.49	±	11
O(7) - C(13)	1.27	±	10
O(8) - C(13)	1.62	±	10
C(13) - C(14)	1.63	±	13
C(14) - C(15)	1.34	±	12
C(15) - C(16)	1.40	±	15
O(9) - C(17)	1.31	±	7
O(10) - C(17)	1.14	±	8

C(17) - C(18)	1.56 ± 10
C(18) - C(19)	1.08 ± 12
C(19) - C(20)	1.74 ± 13
O(11) - C(21)	1.31 ± 9
O(12) - C(21)	1.27 ± 10
C(21) - C(22)	1.40 ± 12
C(22) - C(23)	1.27 ± 12
C(23) - C(24)	1.52 ± 12
O(13) - C(25)	1.16 ± 10
O(14) - C(25)	1.12 ± 10
C(25) - C(26)	1.55 ± 12
C(26) - C(27)	1.48 ± 15
C(27) - C(28)	1.57 ± 18
O(15) - C(29)	1.30 ± 9
O(16) - C(29)	1.49 ± 9
C(29) - C(30)	1.55 ± 10
C(30) - C(31)	1.36 ± 13
C(31) - C(32)	1.73 ± 15

TABLE 6.7

<u>INTERMOLECULAR</u>	<u>O ... O</u>	<u>CONTACTS</u>
O(2) ...	O(16A)	2.62 \pm 7
O(3) ...	O(8B)	2.46 \pm 7
O(6) ...	O(9)	2.38 \pm 6
O(11) ...	O(14C)	2.41 \pm 6

Atoms with coordinates given in Table 6.1 have no extra designation; other symmetry related atoms are distinguished as follows:-

A $x + \frac{1}{2}, y - \frac{1}{2}, 1 - z$

B $x, y, z + 1$

C $x + \frac{1}{2}, y + \frac{1}{2}, 1 - z$

D $x, y, z - 1$

TABLE 6.8RUBIDIUM ... OXYGEN CONTACTS

Rb(1) ... O(12)	2.78
Rb(1) ... O(7B)	2.87
Rb(1) ... O(10C)	2.96
Rb(1) ... O(5)	2.97
Rb(1) ... O(3)	2.98
Rb(2) ... O(15)	2.68
Rb(2) ... O(3D)	2.97
Rb(2) ... O(2C)	3.06
Rb(2) ... O(12)	3.07
Rb(2) ... O(9)	3.09
Rb(2) ... O(5)	3.09
Rb(2) ... O(7)	3.16
Rb(3) ... O(4)	2.76
Rb(3) ... O(6)	2.94
Rb(3) ... O(12A)	2.95
Rb(3) ... O(1)	3.00
Rb(3) ... O(7A)	3.01
Rb(3) ... O(10)	3.18
Rb(4) ... O(1D)	2.75
Rb(4) ... O(4D)	2.79
Rb(4) ... O(11A)	2.91
Rb(4) ... O(8)	2.95

Rb(4) ... 0(6) 3.00

Rb(4) ... 0(10) 3.05

Rb(4) ... 0(13) 3.07

TABLE 6.9

Final observed and calculated structure factors.

M K L F OBS F CALC

M K L F OBS F CALC

M K L F OBS F CALC

M K L F OBS F CALC

M K L F OBS F CALC

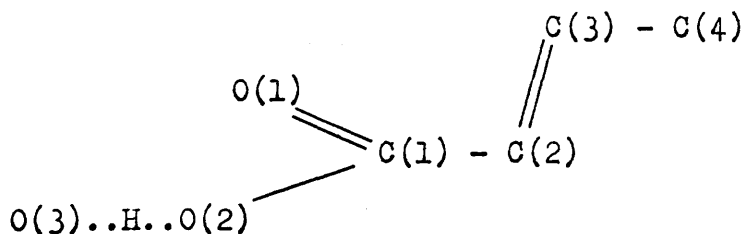
M K L F OBS F CALC

Table with columns M, K, L, F, OBS, F, CALC. The table contains multiple rows of numerical data, organized into several columns. The data appears to be a grid of values, possibly representing a dataset or a calculation table. The columns are labeled with letters and numbers, and the values are numerical. The table is very dense and spans the majority of the page.

M	K	L	F	Obs	F	Calc	M	K	L	F	Obs	F	Calc	M	K	L	F	Obs	F	Calc	M	K	L	F	Obs	F	Calc	M	K	L	F	Obs	F	Calc	M	K	L	F	Obs	F	Calc
6	12	2	71.7	72.6	7	6	3	33.1	31.9	15	0	4	34.3	28.9	4	4	4	12.7	18.4	7	12	4	44.8	37.9	7	6	5	37.3	29.6												
7	12	2	51.0	44.7	13	6	3	65.0	50.5	15	0	4	28.8	28.9	5	8	4	22.3	34.5	11	12	4	37.0	37.2	8	6	5	46.9	41.2												
8	12	2	72.7	73.9	11	7	3	31.8	27.0	21	0	4	34.3	38.0	6	4	7	70.9	74.7	12	12	4	32.1	29.5	10	8	5	24.1	42.9												
9	12	2	42.4	28.3	1	7	3	31.8	27.0	23	0	4	68.4	69.9	9	8	4	181.1	156.4	1	12	4	45.2	44.0	13	6	5	42.9	39.7												
10	12	2	42.7	40.7	2	7	3	30.9	27.0	23	0	4	74.0	69.9	10	8	4	43.9	45.1	2	12	4	41.3	35.9	1	7	5	52.6	46.9												
11	12	2	52.8	53.6	3	7	3	84.4	66.9	25	0	4	55.6	46.4	11	8	4	131.9	133.8	8	12	4	45.4	33.5	2	7	5	51.1	55.0												
12	12	2	30.6	25.4	4	7	3	84.4	66.9	25	0	4	55.6	46.4	11	8	4	131.9	133.8	8	12	4	45.4	33.5	2	7	5	49.6	52.0												
13	12	2	42.8	53.6	5	7	3	30.2	27.0	25	0	4	55.6	46.4	11	8	4	131.9	133.8	8	12	4	45.4	33.5	2	7	5	38.4	29.3												
14	12	2	31.8	32.6	6	7	3	30.2	27.0	25	0	4	18.4	18.3	13	4	4	10.7	45.6	10	13	4	45.4	50.5	9	7	5	27.9	23.2												
0	15	2	42.7	52.4	7	7	3	47.8	44.3	29	0	4	18.7	20.0	15	4	4	59.5	57.2	11	13	4	26.2	29.2	7	7	5	27.9	31.1												
1	13	2	52.1	52.2	8	7	3	36.4	41.9	31	0	4	47.6	25.0	16	4	4	18.8	18.0	12	13	4	37.0	42.2	8	7	5	52.4	52.9												
2	13	2	67.2	60.1	9	7	3	22.3	24.7	31	0	4	51.7	35.0	16	4	4	10.5	105.7	3	1	5	26.1	26.7	10	7	5	52.4	52.9												
3	13	2	52.5	52.6	10	7	3	116.4	105.1	32	0	4	50.3	24.7	16	4	4	24.7	33.3	4	1	5	25.4	22.8	12	7	5	39.2	39.1												
4	13	2	52.9	43.8	11	7	3	42.7	40.4	33	0	4	41.9	24.7	19	4	4	91.2	110.5	8	1	5	17.5	14.7	12	7	5	46.0	39.1												
5	13	2	52.6	57.6	12	7	3	55.6	76.2	35	0	4	44.1	23.4	21	4	4	45.0	49.2	7	1	5	32.8	45.1	15	5	5	31.9	35.9												
6	13	2	43.2	55.2	13	7	3	54.9	46.2	35	0	4	82.9	34.7	25	4	4	31.9	32.0	8	1	5	49.4	40.2	1	8	5	45.2	95.9												
7	13	2	53.1	52.2	14	7	3	28.3	25.7	36	1	4	45.0	46.4	27	4	4	31.1	29.9	10	1	5	48.3	50.5	2	8	5	49.0	54.0												
8	13	2	52.8	61.0	15	7	3	32.1	30.2	37	1	4	47.1	48.0	0	7	4	55.7	53.4	11	1	5	27.0	27.1	3	8	5	51.9	51.5												
9	13	2	30.0	37.0	16	7	3	42.8	39.1	38	1	4	74.6	92.0	1	7	4	12.9	17.5	12	1	5	50.1	71.2	4	8	5	99.2	41.8												
10	13	2	44.1	52.0	17	7	3	76.7	72.3	39	1	4	59.1	85.9	2	7	4	69.1	87.3	13	1	5	45.2	44.2	5	8	5	18.8	15.0												
11	13	2	44.5	37.4	18	7	3	89.3	84.5	40	1	4	17.2	18.6	3	7	4	19.2	15.4	12	1	5	39.3	45.4	7	8	5	59.1	57.4												
3	1	3	34.5	21.7	7	8	3	45.2	42.6	41	1	4	30.3	25.4	5	7	4	22.0	19.9	19	1	5	53.9	59.6	1	9	5	25.9	27.7												
4	1	3	88.6	97.8	8	8	2	45.5	43.2	42	1	4	30.1	30.9	6	7	4	21.9	21.9	20	1	5	21.9	21.9	2	9	5	17.3	14.8												
5	1	3	25.2	24.9	9	8	2	37.1	31.0	24	1	4	31.7	21.2	11	7	4	31.3	26.1	1	2	5	23.0	24.3	4	9	5	52.3	40.1												
6	1	3	27.3	26.6	6	8	3	37.5	22.4	26	1	4	45.4	40.2	12	7	4	129.2	118.5	3	2	5	52.6	45.2	6	9	5	42.4	38.6												
7	1	3	66.7	84.9	7	8	3	90.2	75.5	28	1	4	45.0	30.2	13	7	4	116.2	113.2	4	2	5	32.6	35.2	9	10	5	45.3	34.8												
8	1	3	13.3	23.2	9	8	3	33.2	25.7	29	1	4	25.4	15.8	15	7	4	33.9	40.4	7	2	5	41.5	36.8	6	10	5	31.7	30.7												
9	1	3	68.8	87.5	10	8	3	42.7	41.3	5	2	4	12.7	9.3	16	7	4	98.9	85.1	8	2	5	88.9	62.1	7	10	5	26.0	29.9												
10	1	3	30.2	47.9	13	8	3	74.6	82.3	8	2	4	51.5	50.3	17	7	4	17.9	21.9	8	3	5	56.9	56.9	8	8	5	37.3	42.7												
11	1	3	85.0	89.1	14	8	3	82.6	82.3	9	2	4	82.6	82.3	18	7	4	18.0	24.2	10	4	5	45.4	80.2	10	10	5	37.5	42.7												
12	1	3	102.8	105.4	1	8	3	29.8	29.6	8	2	4	21.5	25.2	20	7	4	10.2	19.3	11	2	5	54.5	61.1	11	10	5	22.8	37.0												
20	1	3	57.6	65.2	4	8	3	52.7	56.3	9	2	4	128.6	147.4	22	7	4	32.1	37.8	12	2	5	44.3	47.5	14	10	5	32.0	48.9												
24	1	3	63.9	51.9	6	8	3	100.0	77.6	10	2	4	39.8	39.6	24	7	4	26.2	26.0	13	2	5	64.0	69.3	15	10	5	50.3	52.7												
1	2	3	21.4	8.7	6	9	3	42.4	36.2	11	2	4	184.8	114.8	28	7	4	52.5	45.7	18	2	5	47.4	46.5	16	10	5	48.5	59.1												
2	2	3	118.9	128.9	14	9	3	45.5	44.4	12	2	4	45.2	43.2	1	8	4	52.9	45.3	17	2	5	24.5	26.3	19	10	5	37.5	36.1												
4	2	3	15.6	4.2	16	9	3	46.9	49.3	15	2	4	44.2	42.9	2	8	4	80.3	73.7	19	2	5	44.3	42.7	19	10	5	25.7	25.7												
5	2	3	171.6	180.7	17	9	3	45.2	43.0	16	2	4	62.7	74.7	4	8	4	44.6	55.4	21	2	5	22.4	29.1	21	10	5	50.2	48.7												
7	2	3	137.8	158.7	2	10	9	41.9	52.2	19	2	4	63.2	65.2	8	8	4	139.2	123.1	23	2	5	38.0	42.8	20	10	5	34.1	34.0												
8	2	3	25.9	31.0	3	10	2	114.9	94.5	21	2	4	52.7	50.5	6	8	4	108.2	102.2	1	3	5	75.5	69.3	4	11	5	22.8	31.1												
9	2	3	36.0	29.1	4	10	2	122.4	116.4	0	3	4	97.5	92.1	8	8	4	31.3	32.7	2	3	5	91.9	45.3	7	11	5	26.7	30.1												
10	2	3	75.2	65.2	5	10	3	42.7	39.1	1	3	4	26.0	16.7	9	8	4	32.8	39.2	3	3	5	52.7	47.0	8	11	5	65.6	57.7												
11	2	3	59.1	59.2	6	10	3	101.9	94.9	2	3	4	71.1	64.9	10	8	4	122.7	120.6	4	3	5	40.9	42.4	11	11	5	71.2	59.3												
12	2	3	137.9	140.5	8	10	3	79.7	22.9	3	3	4	9.0	18.5	12	8	4	17.0	12.2	6	3	5	70.5	68.4	11	11	5	33.0	42.6												
13	2	3	92.8	90.1	11	10	3	98.1	93.1	5	3	4	19.8	19.5	21	8	4	37.2	41.6	7	3	5	12.1	26.5	12	11	5	32.0	28.7												
21	2	3	105.4	105.7	13	10	3	142.7	125.2	6	3	4	74.8	74.8	23	8	4	31.9	29.5	8	3	5	23.4	22.3	14	11	5	22.8	27.0												
22	2	3	96.2	96.1	14	10	3	29.1	29.3	7	3	4	71.1	71.1	8	8	4	151.3	158.9	10	3	5	24.4	27.3	15	11	5	56.8	47.0												
23	2	3	96.2	96.1	14	10	3	29.1	29.3	7	3	4	74.6	74.8	0	9	4	151.3	158.9	10	3	5	24.4	27.3	15	11	5	52.8	47.0												
25	2	3	42.2	44.2	16	10	3	30.2	31.8	9	3	4	51.7	22.0	1	9	4	80.1	97.4	11	3	5	50.5	51.8	16	11	5	48.7	40.2												
2	2	3	91.9	75.9	17	10	3	53.2	59.4	11	3	4	31.7	26.7	2	9	4	126.3	126.0	13	3	5	56.6	66.9	17	11	5	41.2	34.9												
3	2	3	26.8	29.2	18	10	3	31.0	39.3	12	3	4	19.6	24.8	4	9	4	47.4	40.3	15	3	5	33.5	44.7	19	11	5	19.0	19.7												
4	2	3	190.0	185.1	19	10	3	37.4	37.9	13	3	4	30.8	29.9	2	9	4	32.1	31.0	16	3	5	29.7	27.7	20	11	5	46.7	42.4												
5	2	3	25.4	28.4	20	10	3	66.4	56.5	14	3	4	123.2	128.8	8	9	4	31.9	30.4	18	3	5	29.7	27.7	21	11	5	41.2	46.2												
6	2	3	172.1	143.7	23	10	3	53.6	50.5	15	3	4	14.0	8.3	6	9	4	70.2	57.7	18	3																				

TABLE 6.10

AVERAGE BOND-LENGTHS FOR A CROTONATE RESIDUE



Atoms	Mean(Å)	S.D. of Mean (Å)	S.D. (Å)	E.S.D. (av.)(Å)
O(1) - C(1)	1.24	0.02	0.07	0.09
O(2) - C(1)	1.38	0.05	0.14	0.08
C(1) - C(2)	1.51	0.04	0.10	0.11
C(2) - C(3)	1.33	0.04	0.12	0.12
C(3) - C(4)	1.59	0.04	0.12	0.14
O(2)...H...O(3)	2.47	0.06	0.12	0.07

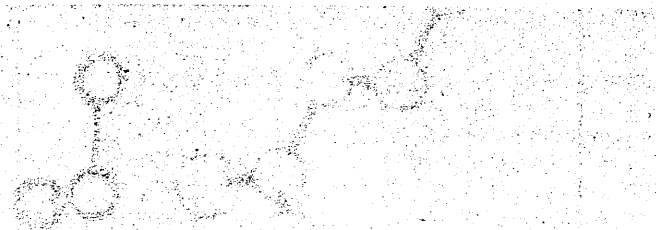
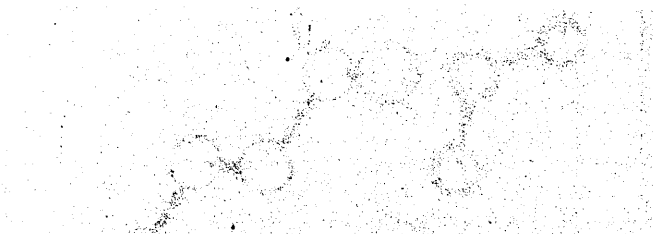
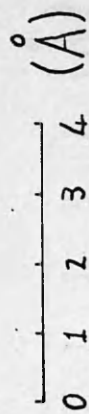
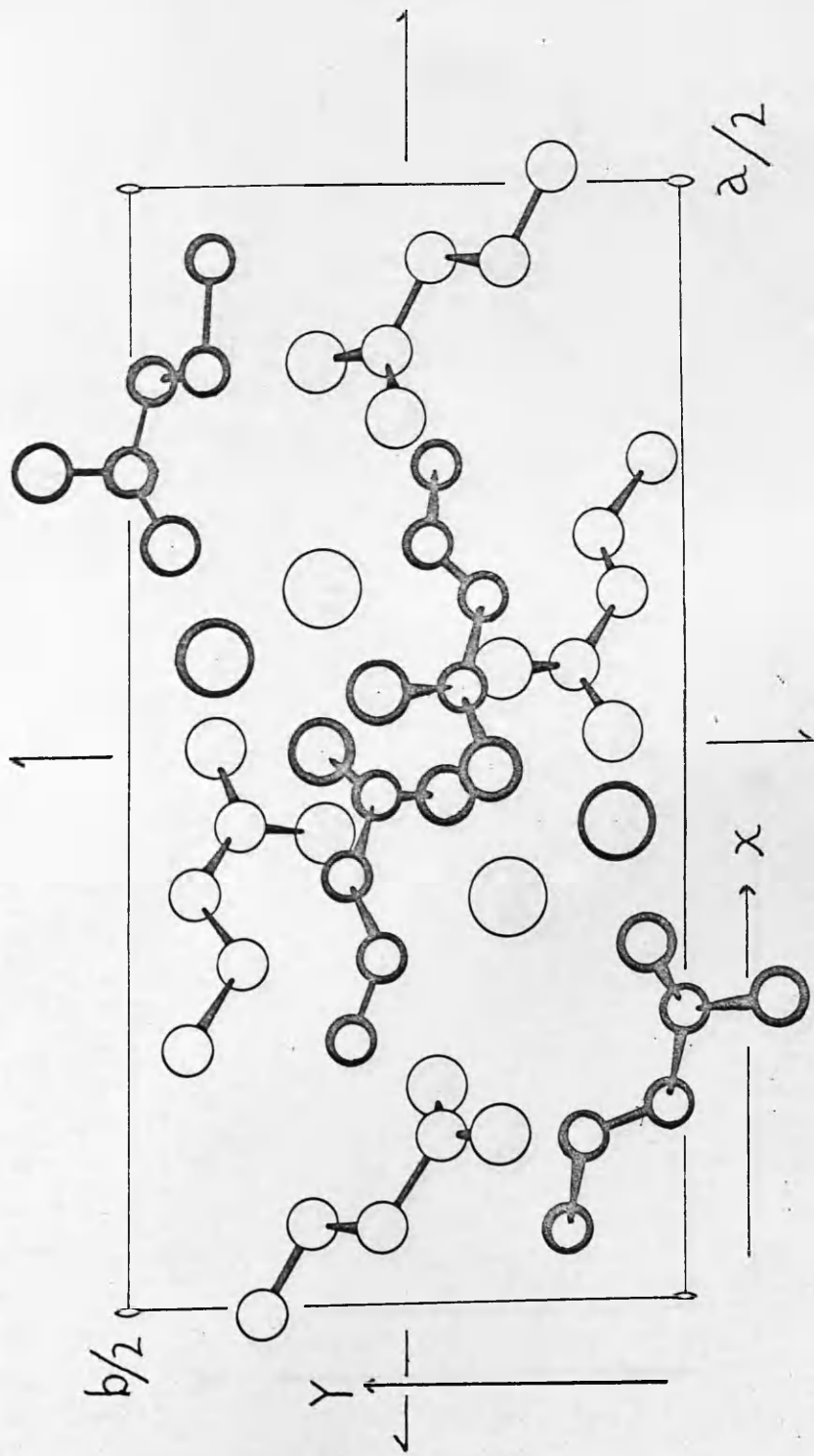
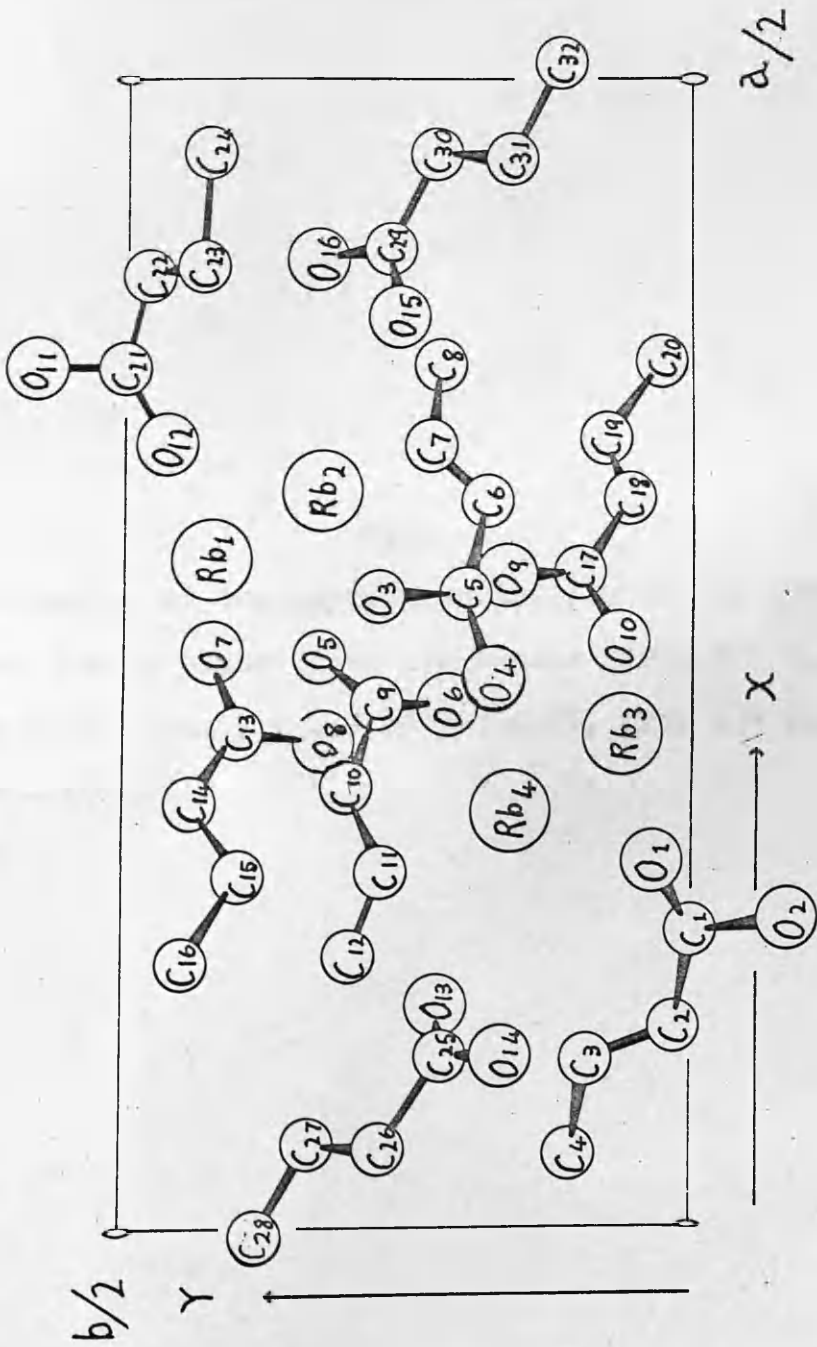


FIG. 1

Projection of the asymmetric portion of the structure onto the xy plane, down the c-axis (from $x = 0.0$ to $x = 0.25$, from $y = 0.0$ to $y = 0.25$, from $z = 0.0$ to $z = 1.0$)



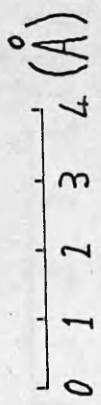
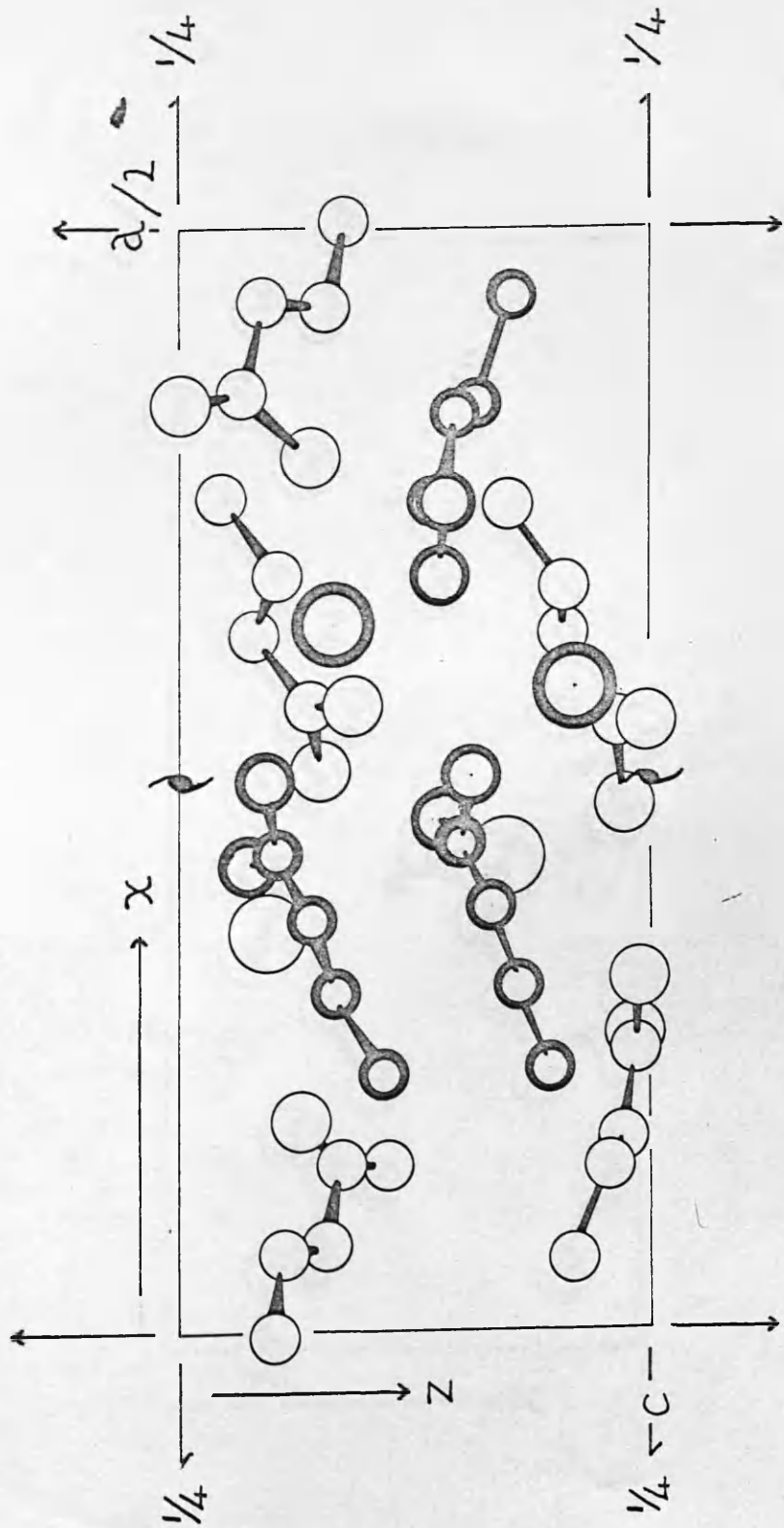




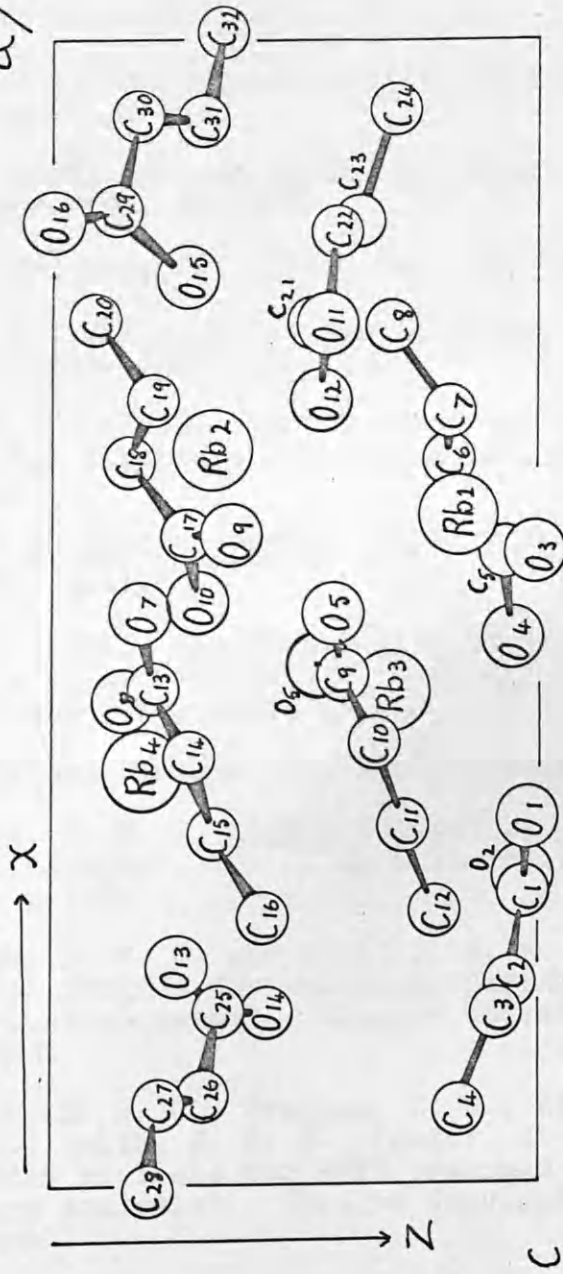
0 1 2 3 4 (Å)

FIG. 2

Projection of the asymmetric portion of the structure onto the xz plane, down the b-axis (from $x = 0.0$ to $x = 0.25$, from $y = 0.0$ to $y = 0.25$, from $z = 0.0$ to $z = 1.0$)



$a/2$



0 1 2 3 4 (Å)

REFERENCES

- Abrahams, S. C. and Speakman, J. C. (1956)
J. Chem. Soc. 2562.
- Arndt. U. W. and Phillips, D. C. (1961) Acta Cryst. 14, 807.
- Beevers, C. A. and Robertson, J. H. (1950)
Acta Cryst. 3, 164.
- Blinč, R. Hadži, D. and Novak, A. (1960)
Z. Electrochem. 64, 567.
- Bohme, H. and Marx, R. (1949) Ber. 74, 1667.
- Bryan, R. F., Mills, H. H. and Speakman, J. C. (1963)
J. Chem. Soc. 4350.
- Buerger, M. J. (1959) "Vector Space and its Application
in Crystal Structure Investigation", John Wiley and Sons,
New York.
- Cheung, K. K. and Sim, G. A. (1964) Nature, Vol. 201,
No. 4925, p.1185.
- Cochran, W. (1966) The Crystalline State, Vol. III,
(Lipson, H. and Cochran, W.)ed. Sir Lawrence Bragg,
G. Bell and Sons, Ltd., London, p.50.
- Cochran, W. and Douglas A.S.(1957) Proc.Roy.Soc.A243, 281
- Cruickshank, D. W. J. (1965) "Computing Methods in
Crystallography", ed. J. S. Rollett, Pergamon Press,
Oxford, p.113.
- Cruickshank, D. W. J. and Smith J. G. F. (1965)
"Computer program for structure refinement by Structure-
Factor-Least-Squares". Glasgow University, Chemistry
Department.
- Cruickshank, D. W. J., Freeman, H. C., Rollett, J. S.,
Sime, J. G., Smith, J. G. F., Truter, M. R., and Wells, M.,
(1964) "Index of lists for KDF9 programs for crystal
structure Analysis". Glasgow University, Chemistry
Department.

- Cruickshank D. W. J. and MacDonald, W. (1967)
Acta Cryst. 22, 37,
- Cruickshank D. W. J. (1965) personal communication.
- Currie, M. Curry, N.A. and Speakman J.C. (1967) to be published.
- Doering, W. von E., and Levy, L. K. (1955)
J. Amer. Chem. Soc. 77, 509.
- Forgie, A., (1964) B.Sc. Thesis, Glasgow University,
Chemistry Department.
- Foster F, and Hargreaves A. (1963) Acta Cryst. 16, 1124-1133.
- Friedel, G. (1913) C. R. Acad. Sci. Paris, 157, 1533.
- Fromm, E. (1889) Ann. 253, 163.
- Geurtz, J. H. Peerdeman, A. F. and Bijvoet, J. M.
(1963) Acta. Cryst. 16, A6.
- Gibson, D. T. (1931) J. Chem. Soc. 2637.
- Golič, L. and Speakman, J. C. (1965A) J. Chem. Soc. 2521.
- Golič, L. and Speakman, J. C. (1965B) J. Chem. Soc. 2530.
- Hall, S. R. and Maslen, E. N. (1965) Acta Cryst. 18, 265.
- Harker, D. and Kasper, J. S. (1948) Acta Cryst. 1, 70.
- Harker, D. (1956), Acta Cryst. 9, 1.
- Harris, D. R. and Mills, H. H. (1966) to be published.
- Hauptman, H. and Karle, J. (1952) Acta Cryst. 5, 60-65.
- Hauptman, H. and Karle, J. (1953) Acta Cryst. 6, 136-141.
- Hibbard, T. N. (1963) Comm. A. C. M. 6, 206.
- Hoogsteen, K. (1957) Ph. D. Thesis, Groningen, Holland.
- International Tables for X-ray Crystallography (1962).
- Islam, I. K. (1967) Structural Work on Phorbol,
Ph.D. Thesis, Glasgow.

- Jacobson, R. A., Wunderlich, J. A. and Lipscomb, W. N.
(1959) Nature, 184, 1719.
- Karle, J. and Hauptman, H. (1953) Acta Cryst. 6, 131-135.
- Karle, I. L. (1961) Acta Cryst. 14, 497.
- Karle, I. L. and Karle, J. (1963) Acta Cryst. 16, 969.
- Karle, I. L., Hauptman, H., Karle, I. and Wing. A. B. (1957)
Acta Cryst. 11, 257.
- Karle, I. L., and Karle, J. (1966A) Acta. Cryst. 21, 860.
- Karle, I. L., and Karle, J. (1966B) Acta Cryst. 21, 555.
- Karle, I. L., and Britts, A. (1966C) Acta Cryst. 21, 532.
- Kendrew, J. C. (1961) Scientific American (December).
- Kipping, F. B. (1935) J. Chem. Soc. 18.
- Kitaigorodskii, A. I., (1965) Acta Cryst. 18, 585.
- McCrorie, I. (1963) B.Sc. Thesis, Glasgow.
- McGregor, D. R. (1966) "The ASS System Mark III", Glasgow
University, Chemistry Department.
- Mills, H. H., and Speakman, J. C. (1961) J. Chem. Soc. 1164.
- Mills, H. H., and Speakman, J. C. (1963) J. Chem. Soc. 4355.
- Muir, K. W. (1967) Ph.d. Thesis, Glasgow University.
- Okaya, Y., Saito, Y. and Papinsky, R., (1953)
Phys. Rev. 98, 1857.
- Phillips, D. C., Blake, C. C. F., Koenig, D. F. Mair, G. A.,
North, A. C. T. and Sarma, V. R. (1965) Nature Vol. 206,
No. 4986, p.263-757.
- Pollard, R. (1966) Structure-factor Least-Squares Program
(isotropic temperature factors), personal communication.
- Ramachandran, G. N. and Raman, S. (1956) Curr. Sci. 25, 384.

- Raman, S. (1958), Proc. Indian Academy of Sciences, XLVII
- Robertson, J. M. (1936) J. Chem. Soc. 1195
- Robertson, J. M. (1963) Proc. Chem. Soc. 229
- Rollett, J. S. (1965) Personal Communication.
- Ross, P. A. (1926), Phys. Rev. 28, 425.
- Samen, E. (1941A), Angew. Chem. 54, 203.
- Samen, E. (1941B), Arkiv. f. Kemi. 14B, No. 28.
- Samen, E. (1942), Arkiv. f. Kemi, 15B, No. 15.
- Samen, E. (1947), Arkiv. f. Kemi, 24B, No. 6.
- Shoemaker, D. P., Donohue, J., Schomaker, V., and Corey, R. B.,
(1950), J. Amer. Chem. Soc. 72, 2328.
- Sayre, D. (1952), Acta. Cryst. 5, 60.
- Silverton, J. V. (1966), Personal Communication.
- Shrivastava, H. N., and Speakman, J. C. (1961)
J. Chem. Soc. 1151.
- Sim, G. A. (1961) Computing Methods and the Phase Problem,
ed. Pepinsky, Robertson and Speakman, Pergamon Press,
Oxford, 235.
- Sim, G. A. (1957) Acta Cryst. 10, 536.
- Sim, G. A. (1959) Acta Cryst. 12, 813.
- Sime, J. G. (1966), Computer Programs for processing
reflexion data from the Linear Diffractometer, and for
Fourier summation. Personal Communication.
- Simpson, P. G., Dobrott, R. D., and Lipscomb, W. N. (1965)
Acta Cryst. 18, 169.
- Skinner, J. M., (1950), Ph.D. Thesis, Glasgow.
- Smith, N. (1949), Ph.D. Thesis, Sheffield.
- Staveley, L. A. K. and Davies, T. (1956) J. Chem. Soc. 2563
- Sutherland, S. A. (1961), Ph.D. Thesis, Glasgow.

Templeton, D. H. (1960) Zeitschrift f. Krist. 113, 234.

Ueki, T., Zalkin, A., and Templeton, D. H. (1966)
Acta Cryst. 20, 836.

Wheatley, P. J. (1954) Acta Cryst. 7, 68.

Woolfson, M. M., (1956), Acta Cryst. 9, 804.

Woolfson, M. M. and Main, P. (1963), Acta Cryst. 16, 1046.

APPENDIX 1

- A.1.1. The Fourier Search Program
- A.1.2. The Sorting Program
- A.1.3. The Weighting-Scheme Program

FOURIER SEARCH PROGRAM→

```

begin
  library A0,A6,A7,A8,A9,A12,A14;
  procedure SORT (n, t, p, in); value n, t, p;
  integer n, t, p; real array in;
  begin
    integer h, i, j, flag;
    real C;
    for j:= 1 step 1 until entier (n/2) do
      begin
        flag:=0;
        for i:= j+1 step 1 until n-j+1 do
          begin
            if in [i,p]> in [i-1,p] then goto L1;
            flag:= 1;
            for h:= 1 step 1 until t do
              begin
                C:=in [i,h]; in [i,h]:= in [i-1,h];
                in [i-1,h]:=C;
              end;
            L1: if in [n-i+1,p]< in [n-i+2,p] then goto L2;
            flag:=1;
            for h:= 1 step 1 until t do
              begin
                C:=in[n-i+1,h]; in [n-i+1,h]:= in [n-i+2,h];
                in [n-i+2,h]:=C;
              end;
            L2:
              end;
            if flag=0 then goto finish;
          end;
        end;
      end;
    finish:
      end procedure SORT;
      procedure SOLVE(n,N,const,coeffts,sing);
      value n; integer sing,n; real array N,const,coeffts;
      begin
        real array L,U[1:n,1:n];
        procedure DECOMP (A,L,U,n,sing); value A,n;
        integer sing,n; real array A,L,U;
        begin
          integer i,j,k;
          for i:= 1 step 1 until n do
            for k:= 1 step 1 until n do
              begin
                if k=i then L[i,k]:=1;
                if k<i then
                  begin
                    if abs(U[k,k])<10-15 then

```

```

    begin
        sing:=0;    goto L4;
    end;
    L[i,k]:=A[i,k]/U[k,k];
    for j:= 1 step 1 until k-1 do
        L[i,k]:=L[i,k]-L[i,j]xU[j,k]/U[k,k];
    U[i,k]:=0;
end;
if k>i then L[i,k]:=0;
if k>i then
begin
    U[i,k]:=A[i,k];
    for j:= 1 step 1 until i-1 do
        U[i,k]:=U[i,k] - L[i,j]xU[j,k];
    end;
end;
L4:
end procedure DECOMP;
procedure INVERT (L,M,n); value L,n;
integer n; real array L,M;
begin
integer i,j,k;
for i:= 1 step 1 until n do
    for k:= 1 step 1 until n do
begin
    if i=k then M[i,k]:=1/L[i,k];
    if i<k then M[i,k]:=0;
    if i>k then
begin
        M[i,k]:=0;
        for j:= k step 1 until i-1 do
            M[i,k]:=M[i,k]-(L[i,j]xM[j,k])/L[i,i];
        end;
    end;
end;
end procedure INVERT;
procedure TRANS (A,B,n); value A,n;
integer n; real array A,B;
begin
integer i,j;
for i:= 1 step 1 until n do
    for j:= 1 step 1 until n do B[i,j]:=A[j,i];
end procedure TRANS;
procedure MM MULT (A,B,n,Y); value A,B,n;
integer n; real array A,B,Y;
begin
integer i,j,k;
for i:= 1 step 1 until n do
    for k:= 1 step 1 until n do

```

```

begin
  Y[i,k]:=0;
  for j:= 1 step 1 until n do
    Y[i,k]:=Y[i,k] + A[i,j]X[j,k];
  end;
end procedure MM MULT;
procedure MV MULT ( A,x,n,b); value A,x,n;
  integer n; real array A,x,b;
begin
  integer i,j;
  for i:= 1 step 1 until n do
    begin
      b[i]:=0;
      for j:= 1 step 1 until n do
        b[i] := b[i] + A[i,j]x[j];
      end;
    end procedure MV MULT;
  sing:=1; DECOMP(N,L,U,n,sing);
  if sing=0 then goto L5;
  INVERT(L,N,n); TRANS(U,L,n); INVERT(L,U,n);
  TRANS(U,L,n); MM MULT(L,N,n,U);
  MV MULT(U,const,n,coeffts);

```

L5:

```

end procedure SOLVE;
integer i, j, k, x, y, z, x min, x max, y min,
  y max, z min, z max, sing,p, Nx, Ny, Nz, l, o, n, q;
real V, a, b, c, s, alpha, beta, gamma, threshold;
open (10); open (20); open (30);
find(100,[DGO30003]); find(101,[20]);
copy text (20, 30, l;l);
begin
  real array unit cell[1:6],S[1:44];
  read binary(100,S,[part2input]);
  read binary(101,unit cell,[unitcell]);
  dataskip(101); Nx:=S[2]; Ny:=S[9]; Nz:=S[16];
  x min:=S[1]; x max:=S[5]; y min:=S[8];
  y max:=S[12]; z min:=S[15]; z max:=S[19];
  a:=unit cell[1]; b:=unit cell[2];
  c:=unit cell[3];

```

alpha:

```

= unit cell[4] x .0174532;
beta:=unit cell[5] x .0174532;
gamma:=unit cell[6] x .0174532;
for i:= 1 step 1 until 3 do
  if abs(S[26+i]-1)<0.0001 then o:=i;
if o=2 then
begin
  o:=3; goto fixed;

```

```

    end;
    if o=3 then o:=2;
next:
    end;
threshold:
=read(20); V:=2xaxbxc/(NxMyNz);
s:=(alpha+beta+gamma)/2;
V:=V*sqrt(sin(s) * sin(s-alpha) * sin(s-beta) *
sin(s-gamma));
if o=2 then
begin
i:=z min; j:=zmax; z min:=y min; zmax:=ymax;
y min:=xmin; y max:=x max; x min:=i; x max:=j;
end;
if o = 3 then
begin
i:=zmin; j:=z max; zmin:=xmin; z max:=xmax;
x min:=y min; xmax:=y max; y min:=i; y max:=j;
end;
p:=q:=n:=0;
begin
real array list[1: 200, 1: 4];
begin
real array density [x min:xmax,y min:y
max], s[1:100,1:3],listt[1:100,1:4];
for z:= z min step 1 until zmax do
begin
l:=0;
read binary (100, density, [Electrondensity]);
for x:= xmin step 1 until xmax do
for y:= ymin step 1 until ymax do
begin
if density[x,y]<threshold then goto continue;
if x+1 >xmax or y+1 >ymax or x-1 <xmin
or y-1 <ymin then goto continue;
for i:= -1 step 1 until 1 do
for j:= -1 step 1 until 1 do
if density[x+i,y+j]>density[x,y] then
goto continue;
l:=l+1; s[l,1]:=x; s[l,2]:=y;
s[l,3]:=density[x,y];
continue:
end;
if l=0 then goto next section;
if n=0 then
begin
for i:= 1 step 1 until 1 do
begin

```

```

list[i,3]:=z; list[i,4]:=s[i,3];
for j:= 1 step 1 until 2 do
  list[i,j]:=s[i,j];
end;
n:=1; goto next section;
end;
for i:= 1 step 1 until 1 do
begin
for j:= 1 step 1 until n do
  if abs(s[i,1]-list[j,1])<1.001 and
  abs(s[i,2]-list[j,2])<1.001 and
  abs(z-list[j,3])<1.001 then
begin
  if s[i,3]>list[j,4] then
  begin
    list[j,1]:=s[i,1]; list[j,2]:=s[i,2];
    list[j,3]:=z; list[j,4]:=s[i,3];
    goto next peak;
  end;
  p:=p+1; listt[p,1]:=s[i,1];
  listt[p,2]:=s[i,2]; listt[p,3]:=z;
  listt[p,4]:=s[i,3];
  if p<40 then goto next peak;
repack: k:=0;
  for x:= 1 step 1 until p do
    if abs(z-listt[x,3])<1.001 then
    begin
      k:=k+1;
      for y:= 1 step 1 until 3 do
        listt[k,y]:=listt[x,y];
      end;
      p:=k; goto next peak;
    end;
  end;
  if p=0 then goto new;
  for j:= 1 step 1 until p do
    if abs(s[i,1]-listt[j,1])<1.001 and
    abs(s[i,2]-listt[j,2])<1.001 and
    abs(z-listt[j,3])<1.001 then
    begin
      if s[i,3]>listt[j,4] then goto new;
      listt[j,1]:=s[i,1]; listt[j,2]:=s[i,2];
      listt[j,3]:=z; listt[j,4]:=s[i,3];
      goto next peak;
    end;
  end;
new: n:=n+1; list[n,1]:=s[i,1];
list[n,2]:=s[i,2]; list[n,3]:=z;
list[n,4]:=s[i,3];
next peak:

```

```

    end;
next section:
    end;
    SORT(n,4,3,list);    l:=0;
    for i:= 1 step 1 until n do
    begin
        list[i,4]:=0;
        if abs(list[i,1]-xmin)>.001 and
           abs(list[i,2]-ymin)>.001 and
           abs(list[i,3]-zmin)>.001 and
           abs(list[i,3]-zmax)>.001 then
        begin
            l:=l+1;
            for j:= 1 step 1 until 3 do
                list[l,j]:=list[i,j];
            end;
        end;
    end;
    rewind(100);
end density and s loop;
begin
    real array values[1:1,-1:1,-1:1,-1:1];
    begin
        real array density[xmin:xmax,ymin:ymax];
        for z:= zmin step 1 until zmax do
        begin
            read binary(100,density,[Electrondensity]);
            for i:= 1 step 1 until l do
            begin
                if list[i,3]-z+1<-.999 then goto NEXT PEAK;
                if list[i,3]-z-1>.999 then
                    goto NEXT SECTION;
                for k:= -1 step 1 until 1 do
                begin
                    if abs(list[i,3]+k-z)<.001 then
                    begin
                        for x:= -1 step 1 until 1 do
                            for y:= -1 step 1 until 1 do
                                begin
                                    values[i,x,y,k]:=
                                        density[list[i,1]+x,list[i,2]+y];
                                    list[i,4]:=list[i,4]+values[i,x,y,k]*V;
                                end;
                            goto NEXT PEAK;
                        end;
                    end;
                end;
            end;
        end;
    end;
NEXT PEAK:
end;
NEXT SECTION:

```



```

    end;
end second density block;
interchange(100);
begin
    real array A[1:27,1:10],const[1:27]
        ,coeffts[1:10],coords[1:3],C[1:3]
        ,LIST[1:8],N[1:10,1:10],vect[1:10],B[1:3,1:3];
    for i:= 1 step 1 until 1 do
        begin
            n:=0;
            for z:= -1 step 1 until 1 do
                for x:= -1 step 1 until 1 do
                    for y:= -1 step 1 until 1 do
                        begin
                            n:=n+1;
                            co: A[n,1]:=1;
                            X: A[n,5]:=list[1,1]+x;
                            Y: A[n,6]:=list[1,2]+y;
                            Z: A[n,7]:=list[1,3]+z;
                            xsq: A[n,2]:=A[n,5]2;
                            ysq: A[n,3]:=A[n,6]2;
                            zsq: A[n,4]:=A[n,7]2;
                            xy: A[n,8]:=A[n,5]xA[n,6];
                            xz: A[n,9]:=A[n,5]xA[n,7];
                            yz: A[n,10]:=A[n,6]xA[n,7];
                            RHS: const[n]:=values[1,x,y,z];
                        end;
                        for j:= 1 step 1 until 10 do
                            begin
                                for k:= 1 step 1 until 10 do N[j,k]:=0;
                                    vect[j]:=0;
                                end;
                                for j:= 1 step 1 until 27 do
                                    for k:= 1 step 1 until 10 do
                                        begin
                                            for n:= k step 1 until 10 do
                                                N[n,k]:=N[k,n]:=N[k,n]+A[j,k]xA[j,n];
                                                vect[k]:=vect[k]+A[j,k]xconst[j];
                                            end;
                                            SOLVE(10,N,vect,coeffts,sing);
                                            if sing=0 then goto L6;
                                            for j:= 1 step 1 until 3 do
                                                B[j,j]:=2xcoeffts[j+1];
                                                B[1,2]:=B[2,1]:=coeffts[8];
                                                B[1,3]:=B[3,1]:=coeffts[9];
                                                B[2,3]:=B[3,2]:=coeffts[10];
                                                C[1]:=-coeffts[5];    C[2]:=-coeffts[6];
                                                C[3]:=-coeffts[7];    SOLVE(3,B,C,coords,sing);

```

```

    if sing=0 then goto L6;
    n:=0;
    for j:= 1 step 1 until 3 do
      if abs(list[i,j]-coords[j])>1.000 then n:=1;
      if n=0 then goto L7;
L6:   for k:= 1 step 1 until 3 do
      coords[k]:=list[i,k];
      list[i,4]:=-999;
L7:   if o =1 then
      for j:= 1 step 1 until 3 do
        LIST[j]:=coords[j];
      if o=2 then
      begin
        LIST[1]:=coords[2];   LIST[2]:=coords[3];
        LIST[3]:=coords[1];
      end;
      if o=3 then
      begin
        LIST[1]:=coords[3];   LIST[2]:=coords[1];
        LIST[3]:=coords[2];
      end;
      LIST[4]:=LIST[1]/Nx;   LIST[5]:=LIST[2]/Ny;
      LIST[6]:=LIST[3]/Nz;   LIST[7]:=list[i,4];
peak density:
      LIST[8]:=coeffts[1]+coeffts[2]xcoords[1]^2
        +coeffts[3]xcoords[2]^2+coeffts[4]
        xcoords[3]^2 +coeffts[5]xcoords[1]
        +coeffts[6]xcoords[2]+coeffts[7]
        xcoords[3]+coeffts[8]xcoords[1]
        xcoords[2]+coeffts[9]xcoords[1]
        xcoords[3]+coeffts[10]xcoords[2]xcoords[3];
      store: q:=q+1;   write binary(100,LIST,[LIST]);
another peak:
      end extrapolation loop;
      end normal equations block;
      end values block;
    end list block;
  interchange(100);   dataskip(100);   skip(100,-q);
  begin
    real array list[1:q,1:8],LIST[1:8];
    for i:= 1 step 1 until q do
      begin
        read binary(100,LIST,[LIST]);
        for j:= 1 step 1 until 8 do list[i,j]:=LIST[j];
      end;
    rewind(100);   interchange(101);
    write text(30,[[2c7s]GRID[10s]COORDINATES[9s]
      FRACTIONAL[2s]COORDINATES[6s]

```

```

    PEAK*WT.***PEAK*HEIGHT[4c]);
    SORT(q,8,7,list);
    for i:= q step -1 until 1 do
    begin
        for j:= 1 step 1 until 3 do
            write(30,format([2s-ndd.ddd]),list[i,j]);
        for j:= 4 step 1 until 6 do
            write(30,format([2s-nd.ddddd]),list[i,j]);
        if abs(list[i,7]+999)<.001 then
            write text(30,[2s]FALSE*PEAK[2c]) else
        begin
            write(30,format([2s-nddd.dd]),list[i,7]);
            write(30,format([2s-nddd.dccc]),list[i,8]);
        end;
    end;
    write text(10,[6c]FRACTIONAL*COORDINATES*FROM*];
    o:=read(20);
    if o=0 then write text(10,[PATTERSON*MAP[4c]);
    if o=1 then write text(10,[Fo*MAP[4c]);
    if o=2 then write text(10,[Fc*MAP[4c]);
    if o=3 then write text(10,[DIFFERENCE*MAP[4c]);
    if o<0 or o>3 then write text(10,[MAP[4c]);
    gap(10,250);
    for i:= q step -1 until 1 do
    begin
        if abs(list[i,7]+999)<.001 then goto last;
        for j:= 4 step 1 until 5 do
            write(10,format([2s-nd.ddddd;]),list[i,j]);
        write(10,format([2s-nd.ddddd;c]),list[i,6]);
        gap(10,15);
    end;
last:
    gap(10,300);
    begin
        real array LL[1:6];
        LL[1]:=q; LL[2]:=8; LL[3]:=o; LL[4]:=Nx;
        LL[5]:=Ny; LL[6]:=Nz;
        write binary(101,LL,[LL]);
        write binary(101,list,[LST]); interchange(101);
        rewind(101);
    end mark block;
    end list and LIST block;
    close(100); close(101); close(30); close(10);
    close(20);
end→

```

CRYSTALLOGRAPHIC DATA INDEX SORTING PROGRAM→

```

begin
  library AO,A6,A7,A8,A9;
  integer t,f,lmin,lmax,n,m,s,np,dv,l,i,j,q,core,
    order,nbl,next,spare,empty;
  integer array ind[-100:100];
  real array LS[1:6];
  open(20); find(101,[20]); find(100,[DG030003]);
  dv:=101; interchange(100); order:=read(20);
  comment order=1 ascending, =2 descending;
  for i:= -100 step 1 until 100 do ind[i]:=0;
  read binary(101,LS,[LS]); f:=read(20); m:=read(20);
  s:=read(20); t:=LS[1]; l:=t+1;
  core:=(6500-52x1)/(t+3); lmax:=-100; lmin:=100;
  begin
    real array block[1:50,1:1],SFS,hold[1:t];
    nbl:=np:=0;
start:
    np:=np+1;
    if np>50 then
      begin
        np:=0; nbl:=nbl+1;
        write binary(100,block,[block]); goto start;
      end;
    read binary(101,SFS,[SFS]);
    if SFS[1]<-998 then
      begin
        block[np,1]:=-999; nbl:=nbl+1;
        block[np,2]:=nbl; np:=0;
        write binary(100,block,[block]);
        interchange(100); rewind(100); rewind(101);
        interchange(dv); write binary(dv,LS,[LS]);
        goto select;
      end;
    for i:= 1 step 1 until t do block[np,i]:=SFS[i];
    block[np,1]:=block[np,s]x10000+block[np,m]
      x100+block[np,f];
    ind[block[np,s]]:=ind[block[np,s]]+1; goto start;
select:
    if order=2 then goto descending order;
    n:=0; lmin:=lmax;
    for i:= -100 step 1 until 100 do
      if ind[i]>0 then
        begin
          m:=ind[i]+n;
          if m>core then goto batch;
          begin
            n:=m; lmax:=i; ind[i]:=0;

```

```

    end;
  end;
  if n=0 then goto finish;
batch:
  begin
    real array in[1:n,1:1],add[1:n,1:2];
    np:=51; q:=0;
  L1: np:=np+1;
    if np>50 then
      begin
        nbl:=nbl-1; read binary(100,block,[block]);
        np:=0; goto L1;
      end;
    if nbl>0 then goto L2;
    if block[np,1]<-998 then
      begin
        nbl:=block[np,2]; rewind(100); goto sorting;
      end;
  L2: j:=block[np,s];
    if j>lmax or j<lmin then goto L1;
    q:=q+1;
    for j:= 1 step 1 until 1 do in[q,j]:=block[np,j];
    add[q,1]:=block[np,1]; add[q,2]:=q; goto L1;
sorting:
  begin
    integer d,k;
    array y[1:2];
    d:=2↑entier(ln(n)/ln(2))-1;
  C1: if d<0 then goto exit;
    i:=1;
  C2: j:=1;
    for k:= 1 step 1 until 2 do y[k]:=add[i+d,k];
  C3: if y[1]<add[j,1] then goto C4;
  C5: for k:= 1 step 1 until 2 do add[j+d,k]:=y[k];
    i:=i+1;
    if i+d<n then goto C2;
    d:=(d-1)/2; goto C1;
  C4: for k:= 1 step 1 until 2 do add[j+d,k]:=add[j,k];
    j:=j-d;
    if j>0 then goto C3;
    goto C5;
  exit:
    end;
rearrange:
  for i:= 1 step 1 until n do
    begin
      if add[i,2]=0 then goto next cycle;
      next:=i; spare:=i;
    end;
  next cycle;

```

```

    for j:= 1 step 1 until t do hold[j]:=in[spare,j];
L3:  empty:=next;  next:=add[empty,2];
    add[empty,2]:=0;
    if next=0 then goto next cycle;
    if next=spare then
    begin
        for j:= 1 step 1 until t do
            in[empty,j]:=hold[j];
        goto next cycle;
    end;
    for j:= 1 step 1 until t do
        in[empty,j]:=in[next,j];
    goto L3;
next cycle:
    for j:= 1 step 1 until t do SFS[j]:=in[i,j];
    write binary(dv,SFS,[SFS]);
    end rearrange block;
    goto select;
    end ascending order block;
descending order:
    n:=0;  lmax:=lmin;
    for i:= 100 step -1 until -100 do
        if ind[i]>0 then
        begin
            m:=ind[i]+n;
            if m>core then goto batch down;
            n:=m;  lmin:=i;  ind[i]:=0;
        end;
        if n=0 then goto finish;
batch down:
    begin
        real array in[1:n,1:1],add[1:n,1:2];
        np:=51;  q:=0;
L10:  np:=np+1;
        if np>50 then
        begin
            nbl:=nbl-1;  read binary(100,block,[block]);
            np:=0;  goto L10;
        end;
        if nbl>0 then goto L20;
        if block[np,1]<-998 then
        begin
            nbl:=block[np,2];  rewind(100);
            goto sorting down;
        end;
L20:  j:=block[np,s];
        if j<lmin or j>lmax then goto L10;
        q:=q+1;

```

```

    for j:= 1 step 1 until 1 do in[q,j]:=block[np,j];
    add[q,1]:=block[np,1];    add[q,2]:=q;    goto L10;
sorting down:
    begin
        integer d,k;
        array y[1:2];
        d:=2↑entier(ln(n)/ln(2))-1;
    C10: if d<0 then goto exit down;
        i:=1;
    C20: j:=i;
        for k:= 1 step 1 until 2 do y[k]:=add[i+d,k];
    C30: if y[1]>add[j,1] then goto C40;
    C50: for k:= 1 step 1 until 2 do add[j+d,k]:=y[k];
        i:=i+1;
        if i+d<n then goto C20;
        d:=(d-1)/2;    goto C10;
    C40: for k:= 1 step 1 until 2 do add[j+d,k]:=add[j,k];
        j:=j-d;
        if j>0 then goto C30;
        goto C50;
exit down:
    end;
rearrange down:
    for i:= 1 step 1 until n do
    begin
        if add[i,2]=0 then goto next cycle down;
        next:=i;    spare:=1;
        for j:= 1 step 1 until t do hold[j]:=in[spare,j];
    L30: empty:=next;    next:=add[empty,2];
        add[empty,2]:=0;
        if next=0 then goto next cycle down;
        if next=spare then
        begin
            for j:= 1 step 1 until t do
                in[empty,j]:=hold[j];
            goto next cycle down;
        end;
        for j:= 1 step 1 until t do
            in[empty,j]:=in[next,j];
        goto L30;
    next cycle down:
        for j:= 1 step 1 until t do SFS[j]:=in[i,j];
        write binary(dv,SFS,[SFS]);
        end rearrange down block;
        goto descending order;
    end descending order block;
finish:
    SFS[1]:=-999;    write binary(dv,SFS,[SFS]);

```

```
interchange(dv); close(101); close(100);  
close(20);  
end;  
end→
```


WEIGHTING LEAST SQUARES DATA PROGRAM→

```

begin
  library A0,A6,A7,A8,A12,A14;
  procedure SOLVE(n,N,const,coeffts,sing);
    value n; integer sing,n; real array N,const,coeffts;
  begin
    real array L,U[1:n,1:n];
    procedure DECOMP(A,L,U,n,sing); value A,n;
      integer sing,n; real array A,L,U;
    begin
      integer i,j,k;
      for i:= 1 step 1 until n do
        for k:= 1 step 1 until n do
          begin
            if k=i then L[i,k]:=1;
            if k<i then
              begin
                if abs(U[k,k])<=10-15 then
                  begin
                    sing:=0; goto I4;
                  end;
                L[i,k]:=A[i,k]/U[k,k];
                for j:= 1 step 1 until k-1 do
                  L[i,k]:=L[i,k]-L[i,j]×U[j,k]/U[k,k];
                U[i,k]:=0;
              end;
            if k>i then L[i,k]:=0;
            if k>i then
              begin
                U[i,k]:=A[i,k];
                for j:= 1 step 1 until i-1 do
                  U[i,k]:=U[i,k]-L[i,j]×U[j,k];
              end;
          end;
        end;
      end;
    I4:
  end procedure DECOMP;
  procedure INVERT(L,M,n); value L,n;
    integer n; real array L,M;
  begin
    integer i,j,k;
    for i:= 1 step 1 until n do
      for k:= 1 step 1 until n do
        begin
          if i=k then M[i,k]:=1/L[i,k];
          if i<k then M[i,k]:=0;
          if i>k then
            begin
              M[i,k]:=0;
            end;
        end;
      end;
    end;
  end;

```

```

    for j:= k step 1 until i-1 do
      M[i,k]:=M[i,k]-(L[i,j]*M[j,k])/L[i,i];
    end;
  end;
end procedure INVERT;
procedure TRANS (A,B,n); value A,n;
  integer n; real array A,B;
begin
  integer i,j;
  for i:= 1 step 1 until n do
    for j:= 1 step 1 until n do B[i,j]:=A[j,i];
  end procedure TRANS;
procedure MM MULT (A,B,n,Y); value A,B,n;
  integer n; real array A,B,Y;
begin
  integer i,j,k;
  for i:= 1 step 1 until n do
    for k:= 1 step 1 until n do
      begin
        Y[i,k]:=0;
        for j:= 1 step 1 until n do
          Y[i,k]:=Y[i,k] + A[i,j]*B[j,k];
        end;
      end procedure MM MULT;
procedure MV MULT ( A,x,n,b); value A,x,n;
  integer n; real array A,x,b;
begin
  integer i,j;
  for i:= 1 step 1 until n do
    begin
      b[i]:=0;
      for j:= 1 step 1 until n do
        b[i] := b[i] + A[i,j]*x[j];
      end;
    end procedure MV MULT;
sing:=1; DECOMP(N,L,U,n,sing);
if sing=0 then goto L5;
INVERT(L,N,n); TRANS(U,L,n); INVERT(L,U,n);
TRANS(U,L,n); MM MULT(L,N,n,U);
MV MULT(U,const,n,coeffts);

```

L5:

```

end procedure SOLVE;
integer i,j,l,sing,p,k,n,no,fc,dc,bc,nobs,put,
  f,g,h,sc,q,r,s;
real fo,del,demax, scale,const,rsinth,sinth;
real array LS[1:6];
f:=format(['-d.dd*nd']); g:=format(['-ndddd.dd']);
q:=format(['-nd;']); r:=format(['ndd.dd;']);

```

```

h:=format([4s-nddd]); s:=format([d.ddw+nd;c]);
open(10); open(20); find(100,[20]);
copy text(20,10,[;]); read binary(100,LS,[LS]);
delmax:=read(20); put:=read(20); nob:=1;
fc:=read(20); dc:=read(20); sc:=read(20);
bc:=read(20); scale:=read(20); p:=read(20)+4;
no:=read(20);

```

begin

```

real array A,vect,coeffts[1:p],fos,nos,sins,
rsins,rsinsq,tot[1:no,1:11],N[1:p,1:p]
,valu[1:no],SFS[1:LS[1]],valu[1:11];

```

for i:= 1 step 1 until p do

begin

vect[i]:=0;

for j:= 1 step 1 until p do N[i,j]:=0;

end;

for i:= 1 step 1 until no do

begin

valu[i]:=read(20);

for j:= 1 step 1 until 11 do

begin

fos[i,j]:=0; sins[i,j]:=0; rsins[i,j]:=0;

tot[i,j]:=0; nos[i,j]:=0; rsinsq[i,j]:=0;

end;

end;

for i:= 1 step 1 until 6 do valu[i+2]:=ix0.1;

valu[1]:=0.01; valu[2]:=0.05; valu[9]:=0.75;

valu[10]:=0.90; valu[11]:=1.5;

start:

read binary(100,SFS,[SFS]);

if SFS[1]<-998 then goto exit;

fo:=SFS[fc]; del:=SFS[dc];

if abs(del)>delmax then

begin

write text(10,[[c]]);

for i:= 1 step 1 until 3 do write(10,g,SFS[i]);

write(10,g,SFS[fc]); write(10,g,SFS[dc]);

write text(10,[[c]]); goto start;

end;

sinth:=SFS[sc]; const:=delT2;

for i:= 1 step 1 until no do if f<valu[i] then

begin

k:=i; goto S1;

end;

S1: for i:= 1 step 1 until 11 do

if sinth<valu[i] then

begin

n:=i; goto S2;

```

S2: tot[k,n]:=tot[k,n]+const; nos[k,n]:=nos[k,n]+1;
fos[k,n]:=fos[k,n]+fo; rsinth:=1/sinth;
rsins[k,n]:=rsins[k,n]+rsinth;
rsinsq[k,n]:=rsinsq[k,n]+rsinth^2;
sins[k,n]:=sins[k,n]+sinth; goto start;
exit:
rewind(100);
for I:= 1 step 1 until no do
  for j:= 1 step 1 until 11 do
    if nos[I,j]>4 or (tot[I,j]>nos[I,j] and
      nos[I,j]>0.001) then
      begin
        del:=nos[I,j]; const:=ln(tot[I,j]/del);
        fo:=fos[I,j]/del;
        for i:= 1 step 1 until p-3 do A[i]:=fo^(i-1);
        A[p]:=sins[I,j]/del; A[p-1]:=rsinsq[I,j]/del;
        A[p-2]:=rsins[I,j]/del;
        for k:= 1 step 1 until p do
          begin
            for n:= k step 1 until p do
              N[k,n]:=N[k,n]+delxA[k]xA[n];
            vect[k]:=vect[k]+delxA[k]xcnst;
          end;
        end;
        for k:= 1 step 1 until p do
          for n:= k step 1 until p do N[n,k]:=N[k,n];
        SOLVE(p,N,vect,coeffts,sing);
        if sing=0 then
          begin
            write text(00,[[c]MATRIX*SINGULAR]); goto L1;
          end;
        gap(10,250);
        for i:= 1 step 1 until no do
          for j:= 1 step 1 until 11 do tot[i,j]:=0;
loop:
read binary(100,SFS,[SFS]);
if SFS[1]<-998 then goto finish;
fo:=SFS[fc]; del:=SFS[dc];
if abs(del)>delmax then goto loop;
sinth:=SFS[sc]; rsinth:=1/sinth;
const:=coeffts[1]+coeffts[p-2]
  xrsinth+coeffts[p-1]xrsinth^2+coeffts[p]xsinth;
for i:= 2 step 1 until p-3 do
  const:=const+coeffts[i]xfo^(i-1);
const:=scalexexp(-const); del:=(del^2)xcnst;
if put=0 then goto noput;
for i:= 1 step 1 until 3 do write(10,q,SFS[i]);

```

```

write(10,q,nobs); write(10,q,SFS[bc]);
write(10,r,fo); write(10,s,sqrt(const));
nopot:
  for i:= 1 step 1 until no do if f<val[i] then
  begin
    j:=i; goto S3;
  end;
S3: for i:= 1 step 1 until 11 do
  if sinh<val[i] then
  begin
    k:=i; goto S4;
  end;
S4: tot[j,k]:=tot[j,k]+del; goto loop;
finish:
gap(10,250);
write text(10,[[2c]Column*batching*by*Fo,
*row*batching*by*sin-theta-by-lambda-all-
squared.[c]First*row*is*average*w*delta*
squared.[c]Second*row*isaverage*Fo.[c]
Third*row*is*average*sin-theta-by-lambda-
all-squared.[c]Fourth*row*is*number*in*
each*batch.[2c]);
for k:= 1 step 1 until no do
begin
write text(10,[[2c]]);
for j:= 1 step 1 until 11 do
if nos[k,j]>0.0001 then
write(10,f,tot[k,j]/nos[k,j]) else
write(10,f,tot[k,j]);
write text(10,[[c]]);
for j:= 1 step 1 until 11 do
if nos[k,j]>0.0001 then
write(10,g,fos[k,j]/nos[k,j]) else
write(10,g,fos[k,j]);
write text(10,[[c]]);
for j:= 1 step 1 until 11 do
if nos[k,j]>0.0001 then
write(10,f,sins[k,j]/nos[k,j]) else
write(10,f,sins[k,j]);
write text(10,[[c]]);
for j:= 1 step 1 until 11 do write(10,h,nos[k,j]);
end;
begin
real array foav,fono[1:no],sinav,sino[1:11];
for i:= 1 step 1 until no do
begin
foav[i]:=0; fono[i]:=0;
end;

```

```

for i:= 1 step 1 until 11 do
  begin
    sinav[i]:=0; sino[i]:=0;
  end;
for i:= 1 step 1 until no do
  for j:= 1 step 1 until 11 do
    begin
      foav[i]:=foav[i]+tot[i,j];
      fono[i]:=fono[i]+nos[i,j];
      sinav[j]:=sinav[j]+tot[i,j];
      sino[j]:=sino[j]+nos[i,j];
    end;
write text(10,[[2c]AVERAGES*BY*FO*AND*
  NUMBERS*IN*BATCHES[c]]);
for i:= 1 step 1 until no do
  if fono[i]>0.001 then
    write(10,f,foav[i]/fono[i]) else
    write(10,f,foav[i]);
write text(10,[[c]]);
for i:= 1 step 1 until no do write(10,h,fono[i]);
write text(10,[[2c]AVERAGES*BY*SINTheta*
  AND*NUMBERS*IN*BATCHES[c]]);
for i:= 1 step 1 until 11 do
  if sino[i]>0.001 then
    write(10,f,sinav[i]/sino[i]) else
    write(10,f,sinav[i]);
write text(10,[[c]]);
for i:= 1 step 1 until 11 do write(10,h,sino[i]);
end;
write text(10,[[2c]Coefficients*of*
  polynomial*used*for*weighting*
  observations.Fo*terms*are*first.[c]
  The*three*sine*theta*terms*are*last.[c]
  They*are*c*coefficients*of*(lambda/sintheta)
  ^2,*(lambda/sintheta)^4,*and*
  (sintheta/lambda)^2.[2c]]);
L1: for j:= 1 step 1 until p do write(10,f,coefrts[j]);
  end;
  close(100); close(10); close(20);
end→

```

APPENDIX 2

A.2. Errors and the Accuracy of Crystal Structure
Structure Determination by X-ray Diffraction

ACCURACY OF STRUCTURES

At the completion of a crystal structure analysis it is usual to calculate a final difference map. This map has features caused by:

1. Random errors in the data,
2. Systematic errors in the data which cannot be accounted for by the adjustment of the model.

If the model can account for a systematic trend in the data then refinement will make it do so, even if this trend is caused by a systematic error. No sign of these errors will appear in a difference map, in the standard deviations or in the R-factor; they will manifest themselves as chemically unsatisfactory models.

Examples of these errors are well known to crystallography:

1. Uncorrected absorption errors will be taken up by adjustments to the temperature factors.
2. Uncorrected anomalous dispersion - in polar space groups - can lead to incorrect atomic positions.
3. Partial site occupancy and structural disorder may be hidden by anisotropic temperature factors.
4. When data have been collected up one axis only, the scale factors and temperature factor components cannot both be refined.

Thus we can see that all our usual criteria, E.S.D., Difference maps and R-factors are powerless to detect this type of error.

As stated previously there are two types of error which are shown up by difference map technique, - random errors and systematic errors which cannot be accounted for by the model.

The magnitude of features on the difference map, - that is, the scale of the topography, - will depend on the average size of the Fourier coefficients used to calculate that map. Thus we can see that it is the absolute size of these coefficients which is important. That is, in a difference map it is the absolute value of Δ which is important, not the relative ratio of Δ to $|F_0|$.

With diffractometer data collected here in Glasgow both the expected errors from counting statistics, and the actual average value of $\langle \overline{\Delta} \rangle$ are smallest for the smallest reflexions (K. W. Muir, 1967). There is absolutely no case for removing them, - in fact they are the most accurate measurements and should in least-squares be given the highest weights.

The R-factor is basically unsuitable therefore as a final criterion of the accuracy of a structure because

the class of small reflexions which have smaller absolute errors, as a class have a very high R-factor.

THE LEAST-SQUARES METHOD

Observations should be given a weight such that the most accurate observations are given the largest weights. Cruickshank has suggested that the average Δ^2 is taken as a measure of the accuracy for reflexions in the same batch by $\sin^2\theta$ or $|F_o|$ magnitude and weights are introduced so that $\langle w \Delta^2 \rangle$ is constant over all the F_o and $\sin^2\theta$ ranges.

When Estimated Standard Deviations (E.S.D.'s) are calculated this technique thus makes allowances both for different accuracies of observation for different sizes of reflexions, and for any systematic errors present in the data which cannot be accommodated by adjustment of the model.

The formula for parameter Estimated Standard Deviations is given below:

$$\sigma_i^2 = (C^{-1})_{ii} - \frac{\sum_{j=1}^m w_j \Delta_j^2}{m - n} ,$$

σ_i = E.S.D. of i^{th} parameter,

$(C^{-1})_{ii}$ = Element of the inverse matrix of the normal equations

m = Number of observations,

n = Number of parameters.

With the coming of the diffractometer two things have happened:

1. m is now much larger since many more observations are taken.
2. The absolute size of Δ has decreased because many of the small reflexions are measured very accurately.

Thus the E.S.D. has become much smaller. In fact it now has reached the point when the uncorrected systematic trends in the data are larger than the uncertainty of measurement.

CONCLUSION

We can no longer expect the calculated E.S.D. to explain away anomalous bond-lengths. The best model may in fact be significantly different from the true structure. Further improvement in accuracy can be obtained only by paying stricter attention to the systematic errors.